Studies in Financial Frontiers: Robo-Advising and Interconnected Markets

by

Jeffrey S. Choy

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Industrial and Operations Engineering)
in the University of Michigan
2024

Doctoral Committee:

Professor Romesh Saigal, Chair
Professor Peter Adriaens
Dr. Abdullah AlShelahi
Professor Larry Seiford

Jeffrey S. Choy
jeffchoy@umich.edu
ORCID iD: 0000-0001-7605-5778

DEDICATION

I dedicate this dissertation to Jess.

ACKNOWLEDGEMENTS

I would firstly like to thank Prof. Romesh Saigal. I am beyond grateful for his mentorship and wisdom over the course of this dissertation. Generous and insightful, he is a pleasure to know and to work with. I would also like to thank Dr. Abdullah AlShelahi, whose advice and collaboration have both been invaluable. For his contributions to my academic and professional development, I express my sincere gratitude. Special thanks to Prof. Peter Adriaens and Prof. Larry Seiford for agreeing to be on my dissertation committee. Many thanks to my Michigan friends for an unforgettable graduate school experience. I am deeply grateful to my parents, whose love and effort supporting all my endeavors cannot be overstated. Finally, I dedicate this dissertation to Jess for her enduring support, patience, and encouragement throughout the years.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF ACRONYMS

DLT Damped Local Trend

GARCH generalized autoregressive conditional heteroskedasticity

GBWM goals-based wealth management

GFC Global Financial Crisis

AUM assets under management

PDE partial differential equation

HJB Hamilton-Jacobi-Bellman

LPPL log-periodic power law

MDP Markov decision process

PPO Proximal Policy Optimization

RL reinforcement learning

ABSTRACT

Financial markets are cornerstones of wealth planning. People turn to these markets to grow their capital and achieve their financial goals. Many have retirement accounts for support at the end of their careers. However, markets are complex, chaotic systems. The interconnected economy means global events can, and do, impact local markets in complicated and surprising fashions. This presents a challenge for long-term investors, for whom important lifestyle tradeoffs must be made for their personal financial goals. To better assist investors in navigating their financial decisions, this dissertation proposes models to not only help them in decision-making for their unique objectives, but also in understanding the impact of external influences on financial markets.

This dissertation contributes to the theory of decision-making in financial markets on two levels: the investor with their unique objectives; and the behavior of the markets in which said investor operates. The first part of this dissertation proposes a continuous-time goals-based wealth management (GBWM) model to maximize the lifetime utility of an investor with multiple competing goals. The model is flexible: the investor has a dedicated retirement account, a market-correlated income, taxes, and consumption considerations. The model consists of a sequence of partial differential equations relating the investor's financial situation at a given time to their optimal portfolio and income allocations. At goal times, the amount contributed to the current goal is optimized, balancing the utility from money withdrawn against expected future utility from money saved. We solve a series of numerical experiments to demonstrate how the investor's optimal decisions vary under different financial circumstances.

The second part looks at GBWM from the perspective of risk management. We examine decision-making in the presence of not only potential recessions, but also uncertainty in an investor's goal times. We present a deep reinforcement learning approach to solve the portfolio and goal contribution problem for a client who invests in a recession-prone economy and whose goals may be random in time. Comparing the recommended portfolio selection and goal contributions reveals how these concerns can be managed practically.

While the first two dissertation components optimize decision-making, the final section examines market abnormalities, the other side of the equation. We develop sensors for external

forces on financial markets at the macroscopic scale. Expanding on the AlShelahi and Saigal (2018) macroscopic model of equity markets, we propose a fluid-dynamical model to characterize market forces, decomposing them into internal and external impacts. We address this by solving a system of stochastic nonlinear partial differential equations, calibrating them with minute-by-minute data from two notable market events: the 2021 GameStop short squeeze and the 2010 flash crash. The results indicate external forces can be detected.

CHAPTER 1

Introduction

Long-term decision-making in financial markets is challenging. The task of matching individual financial goals with sensible investment decisions is a difficult one: not only must the investor's unique financial goals be accounted for, but also the inherent complexities of financial markets. No event better encapsulates the difficulty of long-term investment than the 2007-2008 Global Financial Crisis (GFC), of which the consequences for investors were manifold. Unemployment rose dramatically and household wealth fell [24]. Adjusting to declining incomes, households reduced consumption, increased personal saving rates, and allocated to more conservative portfolios [24, 40]. Retirement behavior also reacted. Surveys indicated many planned on retiring later due to the crisis [40]. More preretirement-age families expressed an unwillingness to take any financial risk, in line with findings that investors' risk aversion increased substantially following the crisis [35, 45].

The GFC affected traditional investment firms' reputations as investment advisors [80]. It prompted discussions on the value of active management [80]. Robo-advising, a class of automated investment services, emerged from the crisis as an attractive new investment vehicle for retail investors [41]. Alongside, there has been considerable development in theory and practice of GBWM, an investment philosophy which tailors financial decisions to the individual and integrates naturally with robo-advising frameworks [41]. Robo-advising is a burgeoning industry. The market is experiencing considerable growth and is responsible for more than one trillion US dollars of assets under management [2].

While GBWM and robo-advising have demonstrated economic viability and consumer demand, the problem of truly personalized automated investment advice remains open. There are, after all, as many investment scenarios as there are investors, but the literature does not yet offer solutions that address the full range of relevant investment factors. In the first component of this dissertation, we present a continuous dynamic programming GBWM model of greater depth than previous models, covering a wider range of important decision-making factors for any investor. Our approach is interpretable: the derivatives of the value functions with respect to the investor's financial situation and goal contributions elucidates the model's recommendations. In addition to a practical solution algorithm for this multidimensional problem, we present a useful lower bound of the investor's value-to-go to contextualize their financial situation.

The second component of this dissertation uses deep reinforcement learning to solve a GBWM problem that, although relatively simple, builds on the use of reinforcement learning in the GBWM literature and complements the features of the continuous dynamic programming model. We examine potential concerns regarding recession risk and goal-time uncertainty in a sequential goal context, optimizing goal contributions and portfolio allocations. The efficacy of both models motivates a combination of their considerations and recommendations within a broader reinforcement learning approach in future.

The bursting of the United States housing bubble during the GFC demonstrates the inseparability of wealth planning and detecting market abnormalities. A similar story is seen in the COVID-19 pandemic: the now-globalized economy is vulnerable to spillover effects in times of crisis [67, 56]. Beyond macroeconomic catalysts of financial disaster, advancements in financial technology have demonstrated the occasional exploitability of markets. The 2010 flash crash is one example: toxic order flow in the high-frequency domain is a plausible explanation for the crash [37]. More recent is the 2021 GameStop short squeeze, in which users of the social media site Reddit instigated a rally in the GameStop stock [61]. These events highlight the effects of external impacts on financial markets and the value of detecting

them for understanding market mechanics.

As described in [4], physics has served an important role in the development of financial and economic theory. Statistical mechanics, in its concern with complex systems of interacting entities, has found utility in a range of financial models [91, 101]. In the third component of this dissertation, our approach is to use statistical mechanics to model external forces on financial markets. Developing on prior work analyzing markets under a macroscopic fluid dynamics description, we develop a system of equations to detect external market forces on the minute-by-minute scale. We validate the system using data from the flash crash and GameStop short squeeze.

1.1 Summary of Contributions

The dissertation is comprised of three manuscripts. Chapter 2 solves the GBWM problem for an investor with multiple goals, a retirement account, and stochastic income, maximizing their expected lifetime utility. We demonstrate that the investor's optimal decisions and expected future utility may be obtained by solving a series of optimal control problems, the solutions which may each be represented as a Hamilton-Jacobi-Bellman (HJB) partial differential equation (PDE). We solve a series of discretized versions of the problem under different parameterizations to demonstrate not only the flexibility of our approach, but also how the optimal decisions are affected by the investor's financial situation.

Chapter 3 solves a GBWM problem for an investor with multiple competing goals in which the goal times are uncertain and the economy is prone to occasional recessions. We demonstrate how the investor's portfolio allocations and goal contributions are affected by these considerations by parameterizing a deep reinforcement learning algorithm with a range of investment scenarios.

In Chapter 4, we expand the model in [5] to include an additional variable, gravity, which we connect to external forces on financial markets. We incorporate the associated stochastic

partial differential equation into the model. For both the 2021 GameStop short squeeze and the 2010 flash crash, we solve the model on minute-by-minute stock data to demonstrate the model's efficacy. We use sentiment data from social media comments during the short squeeze to further validate the model.

All three chapters are presented in a scientific format, each sectioned by the introduction, methods, results, and conclusions. Chapter 5 concludes the dissertation, summarizing the results and suggesting future research directions.

CHAPTER 2

Robo-Advising: Optimal Goals-Based Wealth and Retirement Planning

2.1 Introduction

Robo-advising is a rapidly developing class of online services using algorithms to manage customer investment portfolios and provide financial advice [41]. A client provides relevant information, such as their wealth, investment goals, and risk tolerance. The advisor uses this information, in addition to market analysis, to then synthesize investment recommendations which they automatically implement and manage [39].

The genesis of retail robo-advising, with companies Betterment and Wealthfront, closely followed the 2008 financial crisis [41]. The industry has since experienced rapid growth: Figure 2.1 projects assets under management (AUM) of over US$2 trillion in 2025 [2]. The concomitant growth in the number of robo-advising clients has resulted in approximately 31 million users of robo-advising services as of 2023, more than triple that from even four years ago [2].

The potential benefits of robo-advising are manifold, with perhaps the most attractive being the lower fees in comparison to human advisors. A typical fee for a human advisor may be up to 2% of AUM, while robo-advisors commonly charge much less, at about 0% to 0.5% [23, 41]. Furthermore, robo-advisors can benefit less well-off investors. Minimum balances required to invest can be less of a barrier for robo-advising, as human advisors may

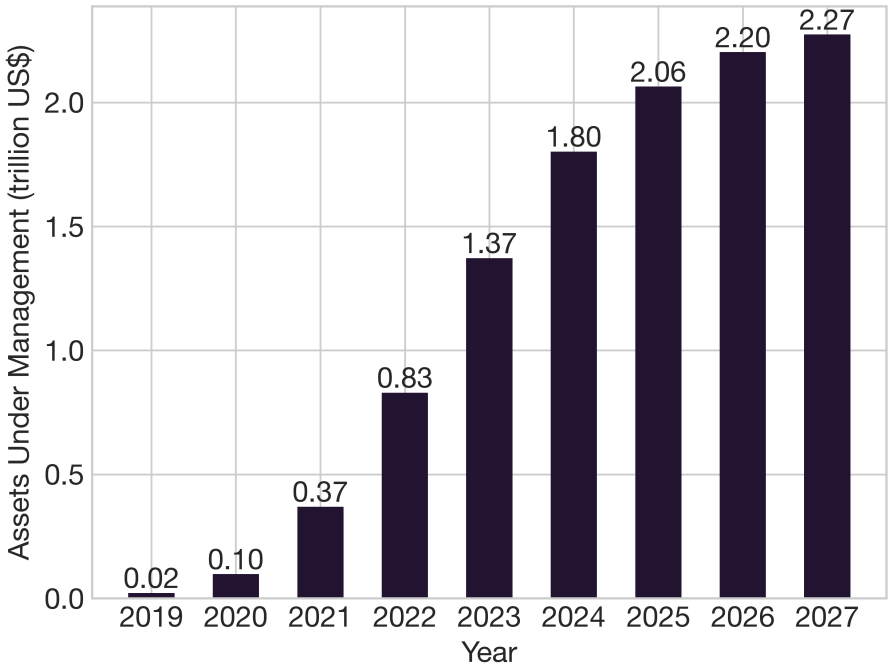

Figure 2.1: Forecasted Assets Under Management of Robo-Advisors. Reproduced from [2].

require a much higher minimum investment for a new client [41].

It is worth noting that, while robo-advising may appear in opposition to human advisors at first glance, various financial firms offer hybrid investment options in which only part of the investment process is automated and the client maintains the ability to consult a financial advisor [41]. Indeed, many benefits of human advisors, such as the ability to empathize with a client's circumstances and assist in improving the client's financial literacy are (currently) beyond robo-advising.

Goals-based wealth management (GBWM) is an investment framework that focuses on helping a client achieve consumption goals. Under GBWM, a client provides their consumption targets, investment horizon, and relative preferences, among other factors, which a portfolio manager uses to construct a strategy to best achieve the client's goals [17]. Fundamentally, GBWM is a behavioral finance undertaking that relies on Kahneman and Tversky's celebrated *prospect theory* which brings psychological considerations to finance [52]. Thaler

introduced the concept of *mental accounting* by which investors mentally partition their portfolios, each with a distinct objective and risk tolerance [86]. Building on these insights, Shefrin and Statman develop *behavioral portfolio theory*, which provides insight on how investors can have different risk-return preferences for different goals. Unifying behavioral finance and modern portfolio theory into a mental accounting (i.e. goal-based portfolio) framework, [26] catalyzed the use of GBWM in the wealth management industry [16, 66].

In this research, we consider a client whose portfolio is managed from the beginning of their career. The client approaches the advisor with a set of consumption goals at particular times, for which they receive some utility for contributing towards. The goals are such that the investor receives utility even if the target consumption is only partially met. Natural examples of such goals include a new car or home for which some, but not all, utility is obtained if less expensive options are purchased. The client also receives utility for immediate consumption of the salary they choose not to invest. In a similar fashion to the consumption goals, the client also has a retirement goal, representing the amount the client would feel comfortable retiring with at the end of their career. In service of this goal, the client has a retirement account separate from their investment portfolio to which they may allocate a proportion of their salary, matched by their employer to some degree. The investment portfolio may be used to satisfy any goal, including the retirement goal. However, the retirement account may be used only to satisfy the retirement goal. The client must, at any time, make two decisions: how to allocate their portfolio funds, and how to allocate their income between their investment portfolio, their retirement account, and immediate consumption. At each of the client's goal times, another decision is made of how much of their portfolio to withdraw towards the current goal at, potentially, the expense of future goals.

Our approach to maximize the investor's expected discounted future lifetime utility involves formulating continuous-time solutions to a series of stochastic dynamic programs. We subsequently calculate solutions to discrete-time approximations of these programs. We use

correlated geometric Brownian motions to represent the risky assets to which the investment portfolio may be allocated, the retirement account asset, and the client's income. At the time of each goal, a convex optimization problem is solved to balance the goal contribution with future expected utility.

We conduct an array of experiments to examine how the optimal investment strategy varies with the client's preferences and circumstances, including the structure of their compensation, their employer retirement contribution, or expected salary growth. In our setting, the client has two intermediate consumption goals in addition to their retirement goal. Our results demonstrate an intuitive set of investor behaviors. Portfolio withdrawals at goal deadlines are affected by the investor's retirement balance and income level, and are sensitive to market-income correlation. Investors also have distinct income allocation decisions leading up to, and directly after, goal deadlines. Risk-aversion close to the investor's target retirement value is reflected in the optimal portfolio and becomes more pronounced closer to retirement.

In addition to introducing an expanded decision-making framework for sequential goals, we contribute an effective, concavity-preserving discrete-time solution approach for our multi-dimensional PDE. We also present a useful lower bound on the investor's value, representing the expected utility of a disengaged investor, to contextualize the value of optimal decision-making.

2.2 Related Works

Our work adds to the GBWM literature by contributing to continuous-time stochastic dynamic programming methods for sequential goals. In comparison with previous studies, we incorporate separate processes for a retirement account and the investor's income, enabling a richer analysis of decision-making pertinent to the investor's lifetime utility. We provide here a summary of relevant literature.

In some cases, particularly with a single consumption goal, explicit solutions for portfolio allocations may be obtained from solving a continuous-time PDE. Examples include Merton's work on portfolio selection under uncertainty, Samuelson's work on lifetime portfolio selection, and Browne's work including forced withdrawal of funds [68, 76, 15]. Other approaches use martingale methods pioneered by Cox and Huang [25]. Wang et al., for example, use a martingale approach to minimize the wealth required for an investor to achieve their goals with a given probability of success [94].

Discrete-time methods for approximation can be used where continuous-time analytical solutions to GBWM problems cannot be found. Das et al. uses a discrete-time dynamic programming algorithm to maximize the probability of an investor attaining their desired wealth level after a certain time [27]. This framework is extended in [28] to account for multiple competing goals. Capponi and Zhang [21] proposes a continous-time method for maximizing the weighted fundedness of a client's goals, conducting a series of comparative statics experiments using a discrete-time approximation.

The rest of the paper is as follows. Section 2.3 details the problem formulation, continuous- and discrete-time solutions, and the parameterizations used for our numerical experiments. The results from, and discussion of, these experiments follow in Section 2.4. Conclusions are given in Section 2.5.

2.3 Methods

The methodology used in this research is structured into four sections. In Section 2.3.1, we specify the problem formulation. We then describe continuous-time solutions to the problem in Section 2.3.2 before detailing a numerical solution procedure for the discrete-time case in Section 2.3.3. Section 2.3.4 presents a detailed description of our numerical experiments.

2.3.1 Formulation

2.3.1.1 The Market and Investor

The system consists of the following components. We have N risky assets with values at time $t \in [0, T]$ of $S(t) = [S_1(t), S_2(t), ..., S_N(t)]^T \in \mathbb{R}_+^N$ into which the portfolio funds, $X_P(t) \in \mathbb{R}_{\geq 0}$, may be allocated with weights $\pi(t) = [\pi_1(t), \pi_2(t), ..., \pi_N(t)]^T \in [0, 1]^N$. Here, we assume the investor adopts a long-only strategy, a natural condition for long-term wealth planning. There also exists a risk-free asset which has the interest rate $r \in \mathbb{R}_{\geq 0}$ into which the investor allocates a proportion of their portfolio value $\pi_B(t) \in [0, 1]$. Therefore, portfolio allocations are constrained so that $\sum_{i=1}^{N} \pi_i(t) + \pi_B(t) = 1$. The investor has a retirement account of value $X_R(t) \in \mathbb{R}_{\geq 0}$, allocated to a retirement fund $S_R(t)$, and a yearly income of $X_I(t) \in \mathbb{R}_{\geq 0}$. Collectively, denote by $X(t) = [X_P(t), X_R(t), X_I(t)]^T$ the vector of financial (or spatial) states at time t.

We now specify how the investor chooses to allocate incoming cash. The investor invests a proportion of their income $\lambda_P(t) \in [0, 1]$ into their portfolio. A proportion of their income is contributed to their retirement account, which is assumed to be employer-matched at a factor of $\kappa \geq 0$. We define this contribution to be $\lambda_R(t) \in [0, \min\{1 - \lambda_P, \gamma\}]$, where $\gamma \in [0, 1]$ is the maximum permissible contribution proportion. The remainder of the income is consumed immediately at a rate of $\lambda_C(t) = 1 - \lambda_P(t) - \lambda_R(t) \in [0, 1]$. Collectively, we denote the above decisions by $\lambda(t) = (\lambda_P(t), \lambda_R(t), \lambda_C(t))$.

We make the following tax assumptions. The investor's income tax is constant at a rate of $\nu_I \in [0, 1]$, which is exacted from portfolio contributions and consumption from income, but not retirement contributions. Portfolio withdrawals are taxed at a rate of $\nu_P \in [0, 1]$ regardless of the portfolio gains or losses. Although the liquidation of a retirement account is typically tax-free, we account for potential retirement taxes in our model, denoted by $\nu_R \in [0, 1]$, which are deducted at the terminal time. We assume that the investor's adjusted gross income is equal to their gross income for simplicity.

The dependence between the processes is specified by the following covariance matrix in $\mathbb{S}_+^{N+2}$:

$$C = \Sigma\Sigma^T \tag{2.1}$$

$$= \left[\begin{array}{c|c|c} C_{PP} & C_{PR} & C_{PI} \\ \hline C_{PR}^T & C_{RR} & C_{RI} \\ \hline C_{PI}^T & C_{RI} & C_{II} \end{array}\right] \tag{2.2}$$

with $C_{PP} \in \mathbb{S}_+^{N\times N}$, $C_{PR} \in \mathbb{R}^N$, $C_{PI} \in \mathbb{R}^N$, and $C_{RR}, C_{II}, C_{RI} \in \mathbb{R}$.

The risky assets, retirement account asset, and income are assumed to follow correlated geometric Brownian motions, with dynamics

$$dS_i(t) = \mu_i S_i(t)dt + S_i(t)\sigma_i d\mathbf{W}(t), \quad i \in \{1, 2, ..., N\} \tag{2.3}$$

$$dS_R(t) = \mu_R S_R(t)dt + S_R(t)\sigma_R d\mathbf{W}(t) \tag{2.4}$$

$$dX_I(t) = \mu_I X_I(t)dt + X_I(t)\sigma_I d\mathbf{W}(t) \tag{2.5}$$

where $\mathbf{W}(t) \in \mathbb{R}^{N+2}$ is an independent multidimensional Brownian motion, $\mu_i > r$, $i = 1, 2, ..., N$ are the drifts of the risky assets, $\mu_R > 0$ and $\mu_I > 0$ are the drifts of the retirement and income processes, respectively, and σ_i is the corresponding row of Σ. We denote by μ the vector of drifts of the risky assets, i.e. $\mu = [\mu_1, \mu_2, ..., \mu_N]^T \in \mathbb{R}^N$, and denote by σ_P the first N rows of Σ. The value of the risk-free asset, denoted by B, follows the dynamics

$$dB(t) = rB(t)\,dt. \tag{2.6}$$

We obtain the following portfolio and retirement account dynamics:

$$dX_P(t) = \left[(1-\nu_I)\lambda_P(t)X_I(t) + \left(r(1-\pi(t)^T\mathbb{1}_N) + \pi(t)^T\mu\right)X_P(t)\right]dt \tag{2.7}$$

$$+ X_P(t)\pi(t)^T\sigma_P d\mathbf{W}(t) \tag{2.8}$$

$$dX_R(t) = \left[(1+\kappa)\lambda_R(t)X_I(t) + \mu_R X_R(t)\right]dt + X_R(t)\sigma_R d\mathbf{W}(t). \tag{2.9}$$

where $\mathbb{1}_N$ is the vector of ones of length N. This yields the following system:

$$\begin{aligned}\begin{bmatrix} dX_P(t) \\ dX_R(t) \\ dX_I(t) \end{bmatrix} &= \begin{bmatrix} r(1-\pi(t)^T\mathbb{1}_N) + \pi(t)^T\mu & 0 & (1-\nu_I)\lambda_P \\ 0 & \mu_R & (1+\kappa)\lambda_R \\ 0 & 0 & \mu_I \end{bmatrix}\begin{bmatrix} X_P(t) \\ X_R(t) \\ X_I(t) \end{bmatrix}dt \\ &+ \begin{bmatrix} X_P(t)\pi(t)^T & 0 & 0 \\ 0 & X_R(t) & 0 \\ 0 & 0 & X_I(t) \end{bmatrix}\Sigma d\mathbf{W}(t).\end{aligned} \tag{2.10}$$

2.3.1.2 Goals

The investor has K goals, the last of which corresponds to their goal retirement wealth. At each non-retirement goal time t_k, for $k = 1, ..., K-1$, the investor chooses to withdraw $\bar{G}_k(X(t_k)) \in [0, X_P(t_k)]$ from their portfolio towards their goal, contributing $G_k(X(t_k)) = (1-\nu_P)\bar{G}_k(X(t_k))$ post-tax. They obtain some utility $u_k(G_k(X(t_k))) \in \mathbb{R}_{\geq 0}$ from this goal. Note that each goal may have a different utility function, enabling different goal priorities and goal-specific utility profiles. For example, the investor may consider a home purchase more important than a car purchase. For the retirement goal at time $t_K = T$, the investor liquidates their portfolio and retirement accounts, thus realizing $(1-\nu_P)X_P(T) + (1-\nu_R)X_R(T)$ in funding and obtaining $u_K((1-\nu_P)X_P(T) + (1-\nu_R)X_R(T))$ in utility. Although there is no goal at time 0, we let $t_0 = 0$ for notational convenience.

We require that each u_k, for $k = 1, ..., K$, is a bounded concave function with $\frac{\partial u_k}{\partial G_k} \to 0$ as $G_k \to \infty$. At time t, the investor receives an instantaneous consumption utility of

$\tilde{u}_k\left((1-\nu_I)\lambda_C(t)X_I(t)\right)$ for some bounded concave utility functions with $\frac{\partial \tilde{u}_k(x)}{\partial x} \to 0$ as $x \to \infty$ for each $k = 1, ..., K$. We also assume that all utilities are discounted continously at rate r and have $u_k(0) = 0$ and $\tilde{u}_k(0) = 0$ for all $k = 1, ..., K$.

2.3.1.3 Policy

We now formalize the client's decision-making process. The goal is to obtain an optimal *policy*, the best mapping from the investor's financial situation to their feasible decisions. In addition to the constraints specified in Section 2.3.1.1, we add a portfolio volatility maximum, $\sigma_{\max}$, to limit the client's risk exposure. We define the convex set of admissible policies across the problem duration

$$\begin{aligned}\mathcal{A} = \big\{(\pi, \lambda_P, \lambda_R, \lambda_C) : \pi^T \mathbb{1}_N \leq 1, \pi \geq 0, \lambda_P \in [0,1], \\ \lambda_R \in [0, \gamma], \lambda_P + \lambda_R + \lambda_C = 1, \pi^T C_{PP} \pi \leq \sigma^2_{\max}\big\}\end{aligned} \tag{2.11}$$

$$\subset \mathbb{R}^{N+3} \tag{2.12}$$

Denoting by $\mathcal{X} = [0, T) \times \mathbb{R}^3_{\geq 0}$ the domain of the investor's financial situation throughout their career, we define a policy to be an element of $\mathcal{F}(\mathcal{X}, \mathcal{A})$, i.e. the set of functions from $\mathcal{X}$ into $\mathcal{A}$, which is itself a subset of the vector space $\mathcal{F}(\mathcal{X}, \mathbb{R}^{N+3})$, having the natural addition and multiplication definitions. It is easy to see that $\mathcal{F}(\mathcal{X}, \mathcal{A})$ is convex under these definitions. For policy $a \in \mathcal{F}(\mathcal{X}, \mathcal{A})$, denote the controls chosen under a by $\pi^a(t, X(t))$, $\lambda^a_P(t, X(t))$, etc. For simplicity, we use $\pi^a(t)$ to refer to $\pi^a(t, X(t))$, and likewise for $\lambda^a_P(t)$, $\lambda^a_R(t)$, and $\lambda^a_C(t)$.

The investor's total expected lifetime utility, which we intend to maximize, can be expressed as

$$\bar{U} = \mathbb{E}^{a^\star}\left[\sum_{k=1}^{K} \underbrace{\int_{t_{k-1}}^{t_k} e^{-rt}\tilde{u}_k((1-\nu_I)\lambda_C^{a^\star}(t)X_I(t))dt}_{\text{discounted consumption utility between } t_{k-1} \text{ and } t_k} + \underbrace{e^{-rt_k}u_k(G_k)}_{\text{discounted goal } k \text{ utility}}\right] \tag{2.13}$$

for the optimal policy, denoted by $a^\star$. Here, $\mathbb{E}^a$ refers to expectation under the policy a.

The following subsection details how the optimal policy $a^\star$ can be characterized via the solution to an Hamilton-Jacobi-Bellman (HJB) partial differential equation (PDE).

2.3.2 Formulating Continuous-Time Solutions

We construct the value function as a representation of the total discounted expected future utility for the investor in any given state under the optimal action. This value function plays a key role in deriving the optimal policy $a^\star$. In the following subsections, we show the value function can be represented as a HJB PDE.

2.3.2.1 The Value Function

We establish a terminal condition, $V_{K+1} = 0$, recognizing that the utility to gain post-retirement is zero. To recursively define the value functions for goals $k = 1, 2, ..., K$, we proceed as follows:

$$V_k(t, X(t)) = \mathbb{E}^{a^\star}\left[U_k(t, X(t)) \mid \mathcal{F}_t\right], \quad t_{k-1} \leq t < t_k \tag{2.14}$$

$$U_k(t, X(t)) = \int_t^{t_k} e^{-r(s-t)} \tilde{u}_k((1-\nu_I)\lambda_C^{a^\star}(s) X_I(s))\, ds + e^{-r(t_k - t)} \Phi_k(X(t_k); V_{k+1}) \tag{2.15}$$

where

$$\Phi_k(X(t_k); V_{k+1}) = \max_{G_k \in [0, (1-\nu_P) X_P(t_k)]} \left[u_k(G_k) + V_{k+1}\left(t_k, \left[X_P(t_k) - \frac{G_k}{1-\nu_P}, X_R(t_k), X_I(t_k)\right]^T\right)\right]. \tag{2.16}$$

and $\mathcal{F}_t$ is the natural filtration generated by $\mathbf{W}$ at time t. Here, Φ_k represents the optimal value from balancing contribution to goal k and the remaining goals. Figure 2.2 below demonstrates, for a problem with three goals, how the value functions are defined over the time domain and how the Φ_k functions serve as terminal conditions for each dynamic

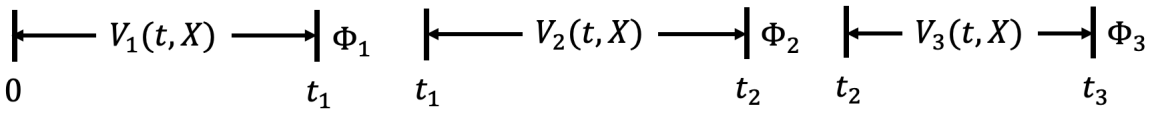

Figure 2.2: Schematic Division of a 3-Goal Time Domain into Separate Value Functions

program.

We have that both V_k and Φ_k, $k = 1, 2, ..., K$ are concave in the financial states as per the following theorem.

Theorem 1. *$V_k(t, x)$ and $\Phi_k(x; V_{k+1})$, $k = 1, 2, ..., K$ are concave in x.*

Proof. We prove the result inductively. Trivially, $V_{K+1} = 0$ is concave. Proving that (V_{k+1} concave $\implies$ Φ_k concave) and (Φ_k concave $\implies$ V_k concave) yields the result.

Suppose V_{k+1} is concave, for some $k = 1, 2, ..., K$. For $j = 1, 2$, let $x^{(j)}$ be arbitrary feasible states at time t_k with respective optimal goal selections $G_k^{(1)}$ and $G_k^{(2)}$ such that, for $\eta \in [0, 1]$, $x^{(3)} \triangleq \eta x^{(1)} + (1-\eta)x^{(2)} \geq 0$. We have that $G_k^{(3)} \triangleq \eta G_k^{(1)} + (1-\eta)G_k^{(2)}$ is a feasible contribution for $x^{(3)}$ as

$$\eta G_k^{(1)} + (1-\eta)G_k^{(2)} \leq (1-\nu_P)(\eta x_P^{(1)} + (1-\eta)x_P^{(2)}) \tag{2.17}$$

$$= (1-\nu_P)x_P^{(3)}. \tag{2.18}$$

Therefore,

$$\begin{aligned}
&\eta\Phi_k(x^{(1)};V_{k+1}) + (1-\eta)\Phi_k(x^{(2)};V_{k+1}) \\
&\quad = \eta u_k(\bar{G}_k^{(1)}) + (1-\eta)u_k(\bar{G}_k^{(2)}) + \eta V_{k+1}\left(t_k, \left[x_P^{(1)} - \frac{\bar{G}_k^{(1)}}{1-\nu_P}, x_R^{(1)}, x_I^{(1)}\right]^T\right) \qquad (2.19)\\
&\qquad + (1-\eta)V_{k+1}\left(t_k, \left[x_P^{(2)} - \frac{\bar{G}_k^{(2)}}{1-\nu_P}, x_R^{(2)}, x_I^{(2)}\right]^T\right) \\
&\quad \leq u_k(\bar{G}_k^{(3)}) + V_{k+1}\left(t_k, \left[x_P^{(3)} - \frac{\bar{G}_k^{(3)}}{1-\nu_P}, x_R^{(3)}, x_I^{(3)}\right]^T\right) \qquad (2.20)\\
&\quad \leq \Phi_k(x^{(3)};V_{k+1}), \qquad (2.21)
\end{aligned}$$

and thus Φ_k is concave. We now look to the value function. For fixed time $t \in [t_{k-1}, t_k)$, suppose we have two feasible financial states, $x^{(1)}$ and $x^{(2)}$. With a slight abuse of notation, let the optimal policies over the duration $[t, t_k)$ for both states $i \in \{1,2\}$ be $a^{(i)} = \left(\pi^{(i)}, \lambda_P^{(i)}, \lambda_R^{(i)}, \lambda_C^{(i)}\right)$. Let, also, $x^{(3)} = \eta x^{(1)} + (1-\eta)x^{(2)}$ for some $\eta \in [0,1]$. Define a new control $a^{(3)} = \left(\pi^{(3)}, \lambda_P^{(3)}, \lambda_R^{(3)}, \lambda_C^{(3)}\right)$ such that

$$\pi^{(3)} = \frac{\eta\pi^{(1)}X_P^{(1)} + (1-\eta)\pi^{(2)}X_P^{(2)}}{\eta X_P^{(1)} + (1-\eta)X_P^{(2)}} \qquad (2.22)$$

$$\lambda_P^{(3)} = \frac{\eta\lambda_P^{(1)}X_I^{(1)} + (1-\eta)\lambda_P^{(2)}X_I^{(2)}}{\eta X_I^{(1)} + (1-\eta)X_I^{(2)}} \qquad (2.23)$$

$$\lambda_R^{(3)} = \frac{\eta\lambda_R^{(1)}X_I^{(1)} + (1-\eta)\lambda_R^{(2)}X_I^{(2)}}{\eta X_I^{(1)} + (1-\eta)X_I^{(2)}} \qquad (2.24)$$

$$\lambda_C^{(3)} = \frac{\eta\lambda_C^{(1)}X_I^{(1)} + (1-\eta)\lambda_C^{(2)}X_I^{(2)}}{\eta X_I^{(1)} + (1-\eta)X_I^{(2)}} \qquad (2.25)$$

where $X^{(i)}$, for $i \in \{1,2,3\}$, is the solution to (2.10) when controlled by $a^{(i)}$ from time t, having started at $x^{(i)}$. It is easy to see that $a^{(3)}$ is a convex combination of $a^{(1)}$ and $a^{(2)}$ and is therefore an admissible control for $X^{(3)}$. We also address the edge cases by requiring $X_P^{(3)} = 0 \implies \pi^{(3)} = 0$ and $X_I^{(3)} = 0 \implies \lambda_P^{(3)} = 1 - \lambda_R^{(3)} = 1 - \lambda_C^{(3)} = 1$. We then have the

dynamics for $i \in \{1, 2\}$ of

$$dX^{(i)}(s) = \begin{bmatrix} r(1 - \pi^{(i)}(s)^T \mathbb{1}_N) + \pi^{(i)}(s)^T \mu & 0 & (1 - \nu_I)\lambda_P^{(i)} \\ 0 & \mu_R & (1 + \kappa)\lambda_R^{(i)} \\ 0 & 0 & \mu_I \end{bmatrix} X^{(i)}(s)dt \tag{2.26}$$

$$+ \begin{bmatrix} X_P^{(i)}(s)\pi^{(i)}(s)^T & 0 & 0 \\ 0 & X_R^{(i)}(s) & 0 \\ 0 & 0 & X_I^{(i)}(s) \end{bmatrix} \Sigma d\mathbf{W}(s) \tag{2.27}$$

over $t \leq s < t_k$. Taking the convex combination of these dynamics, we obtain

$$\begin{aligned} &d(\eta X^{(1)}(s) + (1 - \eta)X^{(2)}(s)) \\ &= \begin{bmatrix} r(1 - \pi^{(3)}(s)^T \mathbb{1}_N) + \pi^{(3)}(s)^T \mu & 0 & (1 - \nu_I)\lambda_P^{(3)} \\ 0 & \mu_R & (1 + \kappa)\lambda_R^{(3)} \\ 0 & 0 & \mu_I \end{bmatrix} X^{(3)}(s)dt \\ &\quad + \begin{bmatrix} X_P(s)^{(3)}\pi^{(3)}(s)^T & 0 & 0 \\ 0 & X_R^{(3)}(s) & 0 \\ 0 & 0 & X_I^{(3)}(s) \end{bmatrix} \Sigma d\mathbf{W}(s) \\ &= dX^{(3)}(s) \end{aligned} \tag{2.28}$$

We therefore have

$$X^{(3)}(s) = \eta X^{(1)}(s) + (1 - \eta)X^{(2)}(s), \quad t \leq s < t_k. \tag{2.29}$$

By the concavity of $\tilde{u}_k$ and Φ_k, we have:

$$U_k(t, X^{(3)}(t)) = \int_t^{t_k} e^{-r(s-t)} \tilde{u}_k((1-\nu_I)\lambda_C^{(3)}(s) X_I^{(3)}(s))\, ds + e^{-r(t_k-t)} \Phi_k(X^{(3)}(t_k); V_{k+1})$$

(2.30)

$$= \int_t^{t_k} e^{-r(s-t)} \tilde{u}_k(\eta(1-\nu_I)\lambda_C^{(1)}(s) X_I^{(1)}(s) + (1-\eta)(1-\nu_I)\lambda_C^{(2)}(s) X_I^{(2)}(s))\, ds$$
$$+ e^{-r(t_k-t)} \Phi_k(\eta X^{(1)}(t_k) + (1-\eta) X^{(2)}(t_k); V_{k+1})$$

(2.31)

$$\geq \eta \left[\int_t^{t_k} e^{-r(s-t)} \tilde{u}_k((1-\nu_I)\lambda_C^{(1)}(s) X_I^{(1)}(s))\, ds + e^{-r(t_k-t)} \Phi_k(X^{(1)}(t_k); V_{k+1}) \right]$$
$$+ (1-\eta) \left[\int_t^{t_k} e^{-r(s-t)} \tilde{u}_k((1-\nu_I)\lambda_C^{(2)}(s) X_I^{(2)}(s))\, ds + e^{-r(t_k-t)} \Phi_k(X^{(2)}(t_k); V_{k+1}) \right]$$

(2.32)

$$= \eta U(t, X^{(1)}(t)) + (1-\eta) U(t, X^{(2)}(t)) \tag{2.33}$$

As $a^{(1)}$ and $a^{(2)}$ are optimal, we have

$$V(t, X^{(3)}(t)) \geq \eta V(t, X^{(1)}(t)) + (1-\eta) V(t, X^{(2)}(t)) \tag{2.34}$$

and thus the value function is concave.

□

As time t determines the upcoming goal k, we denote

$$V(t, X) \triangleq V_{\tilde{k}(t)}(t, X) \tag{2.35}$$

$$\tilde{k}(t) = \arg\min_k \{t < t_k\} \tag{2.36}$$

for notational simplicity.

The existence of a supremum of the value function is guaranteed as all utility functions

are bounded above. This may be expressed as

$$\begin{aligned}\sup_X V(t,X) = \sup_z \tilde{u}_{\tilde{k}(t)}(z) \int_t^{t_k} e^{-r(s-t)}\, ds + e^{-r(t_k-t)} \sup_z u_{\tilde{k}(t)}(z) \\ + \sum_{j=\tilde{k}(t)+1}^{K} \left(\sup_z \tilde{u}_j(z) \int_{t_{j-1}}^{t_j} e^{-r(s-t)}\, ds + e^{-r(t_j-t)} \sup_z u_j(z) \right)\end{aligned} \tag{2.37}$$

We may express the quality of each state relative to the best-case scenario with

$$\zeta(t,X) = \frac{V(t,X)}{\sup_{X'} V(t,X')} \in [0,1], \quad t \in [0,T) \tag{2.38}$$

This quantifies what proportion of the potential discounted future utility the investor is expected to attain under optimal behavior.

The following subsection details how the HJB PDE is formulated from the value function and subsequently solved.

2.3.2.2 Constructing the Hamilton-Jacobi-Bellman Partial Differential Equations

Our problem-solving approach is outlined as follows. Suppose, for problem k, we have solved all problems from $k+1$ to K. In this case, we have knowledge of $\Phi_k(X(t_k))$ and can therefore determine the optimal goal contribution, denoted by $G_k^\star$.

Assumption 1. *There exists a series of C^2 functions $F_k : \mathbb{R}^4 \to \mathbb{R}$, $k = 1, 2, ..., K$ such that*

$$F_k(t, X(t)) = V_k(t, X(t)), \quad t_{k-1} \leq t < t_k \tag{2.39}$$

with terminal condition

$$F_k(t_k, X(t_k)) = \Phi_k(X(t_k); F_{k+1}) \tag{2.40}$$

We then have the following theorem.

Theorem 2. *F_k, $k = 1, 2, ..., K$ as described satisfy the HJB equations*

$$\begin{aligned} rF_k(t, X(t)) = &\frac{\partial F_k}{\partial t} + \frac{\partial F_k}{\partial X_R(t)}(\mu_R X_R(t)) + \frac{\partial F_k}{\partial X_I(t)}(\mu_I X_I(t)) \\ &+ \frac{1}{2}\frac{\partial^2 F_k}{\partial X_R^2(t)} X_R(t)^2 C_{RR} + \frac{1}{2}\frac{\partial^2 F_k}{\partial X_I^2(t)} X_I(t)^2 C_{II} + \frac{\partial^2 F_k}{\partial X_R(t)\partial X_I(t)} X_R(t) X_I(t) C_{RI} \\ &+ \max_{a \in \mathcal{F}(\mathcal{X}, \mathcal{A})} H(t, X(t); a) \end{aligned} \tag{HJB-PDE}$$

where

$$\begin{aligned} H(t, X(t); a) = &\tilde{u}_k((1 - \nu_I)\lambda_C^a(t) X_I(t)) \\ &+ \frac{\partial F_k}{\partial X_P(t)}((r(1 - \pi^{a\,T}(t)\mathbb{1}_N) + \pi^{a\,T}(t)\mu) X_P(t) + (1 - \nu_I)\lambda_P^a(t) X_I(t)) \\ &+ \frac{\partial F_k}{\partial X_R(t)}((1 + \kappa)\lambda_R^a(t) X_I(t)) + \frac{1}{2}\frac{\partial^2 F_k}{\partial X_P^2(t)} X_P^2(t) \pi^{a\,T}(t) C_{PP} \pi^a(t) \\ &+ \frac{\partial^2 F_k}{\partial X_P(t)\partial X_R(t)} X_P(t) X_R(t) \pi^{a\,T}(t) C_{PR} + \frac{\partial^2 F_k}{\partial X_P(t)\partial X_I(t)} X_P(t) X_I(t) \pi^{a\,T} C_{PI} \end{aligned} \tag{2.41}$$

Proof. For $k = 1, 2, ..., K$, applying Ito's lemma to F_k results in

$$\begin{aligned} &dF_k(t, X(t); a^\star) \\ &= \left[\frac{\partial F_k}{\partial t} + \frac{\partial F_k}{\partial X_P(t)}((r(1 - \pi^{a^\star}(t)^T \mathbb{1}_N) + \pi^{a^\star}(t)^T \mu) X_P(t) + (1 - \nu_I)\lambda_P^{a^\star}(t) X_I(t)) \right. \\ &+ \frac{\partial F_k}{\partial X_R(t)}(\mu_R X_R(t) + (1 + \kappa)\lambda_R^{a^\star}(t) X_I(t)) + \frac{\partial F_k}{\partial X_I(t)}(\mu_I X_I(t)) + \frac{1}{2}\frac{\partial^2 F_k}{\partial X_P^2(t)} X_P(t)^2 \pi^{a^\star}(t)^T C_{PP} \pi^{a^\star}(t) \\ &+ \frac{1}{2}\frac{\partial^2 F_k}{\partial X_R^2(t)} X_R(t)^2 C_{RR} + \frac{1}{2}\frac{\partial^2 F_k}{\partial X_I^2(t)} X_I(t)^2 C_{II} + \frac{\partial^2 F_k}{\partial X_P(t)\partial X_R(t)} X_P(t) X_R(t) \pi^{a^\star}(t)^T C_{PR} \\ &\left. + \frac{\partial^2 F_k}{\partial X_P(t)\partial X_I(t)} X_P(t) X_I(t) \pi^{a^\star}(t)^T C_{PI} + \frac{\partial^2 F_k}{\partial X_R(t)\partial X_I(t)} X_R(t) X_I(t) C_{RI} \right] dt \\ &+ \begin{bmatrix} X_P(t)\pi^{a^\star}(t)^T & 0 & 0 \\ 0 & X_R(t) & 0 \\ 0 & 0 & X_I(t) \end{bmatrix} \Sigma d\mathbf{W}(t) \end{aligned}$$

The recursive formulation of the value function obtains

$$F_k(t, X(t)) = \mathbb{E}^{a^\star}\left[\int_t^{\bar{t}} e^{-r(s-t)}\tilde{u}_k((1-\nu_I)\lambda_C^{a^\star}(s)X_I(s))ds + e^{-r(\bar{t}-t)}F_k(\bar{t}, X(\bar{t}))\right] \tag{2.42}$$

for $t_{k-1} \leq t < \bar{t} < t_k$ and $k = K, ..., 1$. Multiplying the value function by e^{-rt} and taking its stochastic-differential form, we obtain the HJB formulation:

$$\mathbb{E}^{a^\star}\left[e^{-rt}\tilde{u}_k((1-\nu_I)\lambda_C^{a^\star}(t)X_I(t))dt + d\left(e^{-rt}F_k(t, X(t))\right)\right] = 0 \tag{2.43}$$

$$\implies \mathbb{E}^{a^\star}\left[\tilde{u}_k((1-\nu_I)\lambda_C^{a^\star}(t)X_I(t))dt + dF_k(t, X(t))\right] = rF_k(t, X(t))dt \tag{2.44}$$

for $t_{k-1} \leq t < t_k$ and $k = K, ..., 1$. As the value function under optimal action is a martingale, we obtain (HJB-PDE) [30]. □

The following subsection details the methods used for solving this problem in discrete time.

2.3.3 Solving the Discrete-Time Case

In the absence of closed-form solutions to (HJB-PDE), a discrete-time solution to the problem is required. The discrete-time approximation to the continuous-time solution takes the form of a series of dynamic programs, each corresponding to a specific goal, which are solved in reverse order. This involves solving for the optimal action and value function at uniformly spaced times, progressing backward from retirement.

Each problem is solved with policy iteration. Our approach uses a Chebyshev polynomial approximation to the value function at each such time. This is motivated and described in the following subsection. In attempting to solve the program, a finite difference approximation of the value function is also tried, unsuccessfully. This approach fails to maintain numerical concavity of the value function. For completeness, the details of this implementation are provided in Appendix A.1.

2.3.3.1 Expanded Chebyshev Polynomial Approximation of the Value Function

Motivation Issues maintaining value function concavity in finite difference approaches motivate the use of methods that guarantee concavity throughout the computation. *Shape-preserving dynamic programming* is a framework to approximate the value function via an interpolation that simultaneously enforces monotonicity and concavity [20]. Chebyshev polynomials may be used as the basis for one such interpolation, and have been applied to economic and financial optimal control problems [44].

Chebyshev Polynomials A polynomial approximation of F_k, $k = 1, 2, ..., K$, is motivated by the Stone-Weierstrass Theorem, reproduced from [95]:

Theorem 3. *(Stone-Weierstrass Theorem) If X is any compact space, let A be a subalgebra of the algebra $C(X)$ over the reals $\mathbb{R}$ with binary operations $+$ and $\times$. Then, if A contains the constant functions and separates the points of X (i.e., for any two distinct points x and y of X, there is some function f in A such that $f(x) \neq f(y)$), A is dense in $C(X)$ equipped with the uniform norm.*

The set of Chebyshev polynomials defined over the problem's spatial domain, expressed in (2.56) below, provides such a subalgebra. The theorem thus motivates the use of Chebyshev polynomial approximations of F_k, for $k = 1, 2, ..., K$.

We now describe our approach to using Chebyshev polynomials for approximation (and interpolation) of the value function over the spatial states. The Chebyshev polynomials for $z \in [-1, 1]$ are defined as follows:

$$T_0(z) = 1 \tag{2.45}$$

$$T_1(z) = z \tag{2.46}$$

$$T_{j+1}(z) = 2xT_j(z) - T_{j-1}(z), \quad j = 1, 2, ... \tag{2.47}$$

Figure 2.3 below visualizes the first five Chebyshev polynomials.

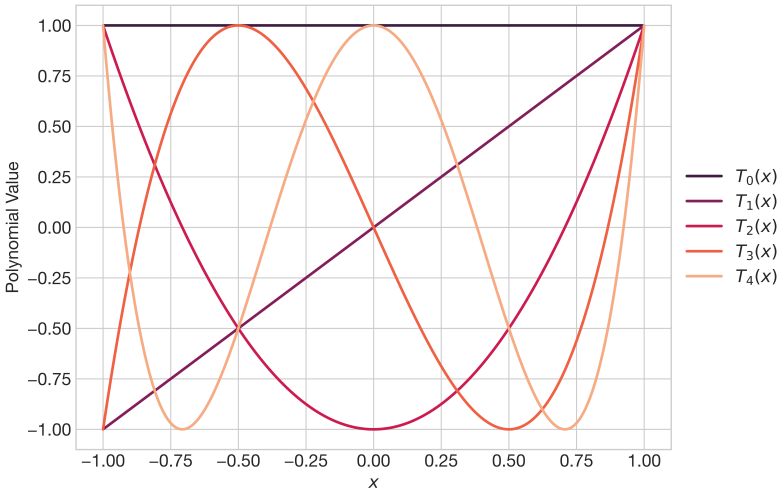

Figure 2.3: The First Five Chebyshev Polynomials

We use a set of Chebyshev nodes scaled over the spatial domain $[0, P_{\max}] \times [0, R_{\max}] \times [0, I_{\max}]$, for some $(P_{\max}, R_{\max}, I_{\max})$ sufficiently large to capture the range of investor circumstance. For example, if u_K admits a maximum, $R_{\max}$ should be no less than $\min\{x : u_K(x) = \max_{x'} u_K(x')\}$ to reveal how investor choices vary as their retirement portfolio value approaches their target. The time domain is discretized uniformly with spacing Δ_t such that $(t_k \mod \Delta_t) = 0$ for all $k = 1, 2, ..., K$. This ensures a value function estimation is guaranteed at each goal time, allowing for optimization of the goal contributions without interpolation in the time dimension. We partition the time steps into each program with

$$\mathbf{T} = \{0, \Delta_t, 2\Delta_t, ..., t_K\} \tag{2.48}$$

$$\mathbf{T}_k = \{t \in \mathbf{T} \mid t_{k-1} \leq t < t_k\}, \quad k = 1, ..., K \tag{2.49}$$

With a slight abuse of notation, denote by $\mathbf{D} = \{P, R, I\}$ the set of spatial dimensions of the problem. For dimension $D \in \mathbf{D}$, let $m_D \in \mathbb{N}$ denote the number of Chebyshev nodes

used for that dimension. Now, for $D \in \mathbf{D}$, let

$$\boldsymbol{z}_D = \left[z_D^{(1)}, ..., z_D^{(m_D)}\right]^T \in [-1, 1]^{m_D}, \tag{2.50}$$

$$z_D^{(i)} = -\cos\left(\frac{(2i-1)\pi}{2m_D}\right), \quad i = 1, ..., m_D \tag{2.51}$$

represent the Chebyshev nodes for each dimension. Scaling these to their respective domain limits, we define, for $D \in \mathbf{D}$,

$$\boldsymbol{y}_D = \left[y_D^{(1)}, ..., y_D^{(m_D)}\right]^T \in [0, D_{\max}]^{m_D} \tag{2.52}$$

$$y_D^{(i)} = \frac{(z_D^{(i)} + 1)(D_{\max} + 2\delta_D)}{2} - \delta_D \in [0, D_{\max}], \quad i = 1, ..., m_D \tag{2.53}$$

$$\delta_D = \frac{z_D^{(1)} + 1}{-2z_D^{(1)}} D_{\max} \tag{2.54}$$

thus ensuring that $y_D^{(1)} = 0$ and $y_D^{(m_D)} = D_{\max}$. This construction yields the *expanded* Chebyshev polynomial approximation, which has superior approximation qualities near the end points [19]. Figure 2.4 below shows the positions of the Chebyshev nodes, i.e. $\{(y_1, y_2, y_3) \mid y_1 \in \boldsymbol{y}_P, y_2 \in \boldsymbol{y}_R, y_3 \in \boldsymbol{y}_I\}$, for $m_P = m_R = m_I = 30$, colored by the terminal value function for the "baseline" investor described in Section 2.3.4.1.

For $D \in \mathbf{D}$, let

$$T_j^D(y) = T_j\left(\frac{2(y + \delta_D)}{D_{\max} + 2\delta_D} - 1\right), \quad j = 0, 1, ... \tag{2.55}$$

representing the Chebyshev polynomial evaluated after mapping state value $y \in [0, D_{\max}]$ to its corresponding value in $[-1, 1]$.

Denote by $d_D \in \mathbb{N}$ the dimension of the polynomial for state $D \in \mathbf{D}$. A tensor product of the univariate Chebyshev polynomials is taken to form a tensor product basis for functions of the spatial dimensions, allowing the construction of the set of the Chebyshev polynomials

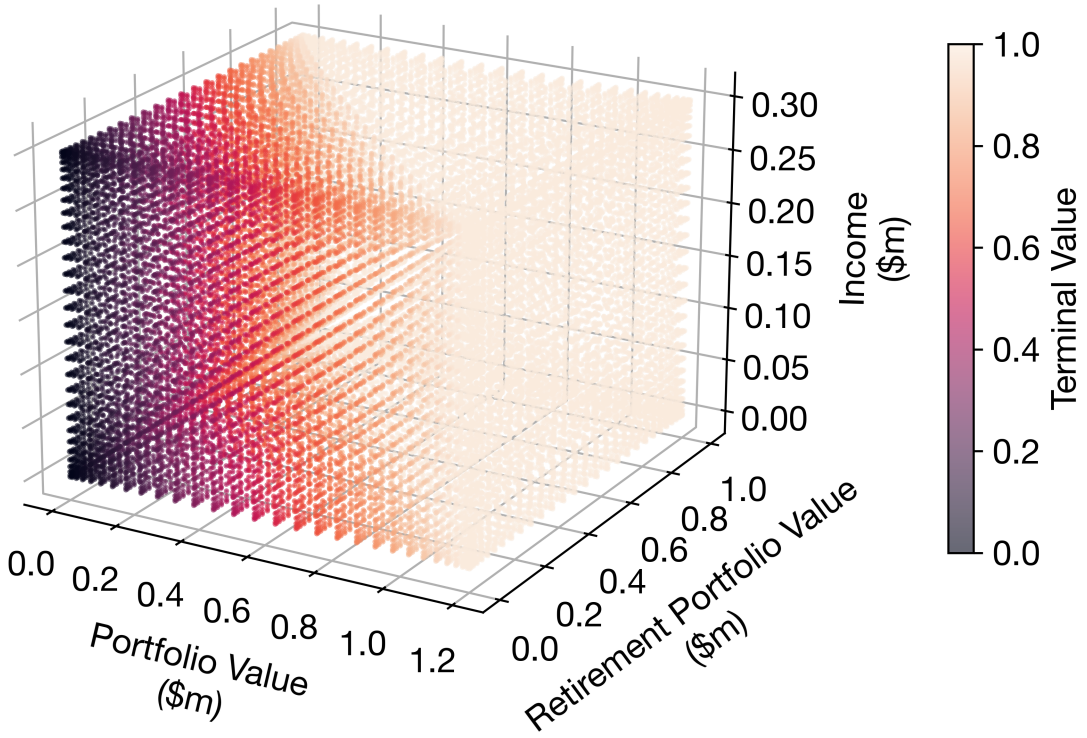

Figure 2.4: Terminal Value Function at the Chebyshev Nodes for the Baseline Investor

over the spatial dimensions:

$$A = \Bigg\{ \sum_{i=0}^{d_P} \sum_{j=0}^{d_R} \sum_{l=0}^{d_I} c_{j_P j_R j_I} T_{j_P}^P (X_P) \, T_{j_R}^R (X_R) \, T_{j_I}^I (X_I) \;\Big|\; c_{j_P j_R j_I} \in \mathbb{R}, \\ j_P = 0, 1, ..., d_P;\; j_R = 0, 1, ..., d_R;\; j_I = 0, 1, ..., d_I \Bigg\} \tag{2.56}$$

Our value function approximation for goal $k = 1, 2, ..., K$ and time $t \in \mathbf{T}_k$ can therefore be parameterized by a set of coefficients $\mathbf{c}^t \in \mathbb{R}^{d_P \times d_R \times d_I}$ as follows:

$$\hat{F}_k(t, X; \mathbf{c}^t) = \sum_{i=0}^{d_P} \sum_{j=0}^{d_R} \sum_{l=0}^{d_I} c_{j_P j_R j_I}^t T_{j_P}^P (X_P) \, T_{j_R}^R (X_R) \, T_{j_I}^I (X_I) \quad t \in \mathbf{T}_k \tag{2.57}$$

$$\mathbf{c}^t = \left\{ c_{j_P j_R j_I}^t \right\}_{i=0,1,...,d_P;\, j=0,1,...,d_R;\, l=0,1,...,d_I} \tag{2.58}$$

where $\hat{F}_k$ is the value function approximation for goal k.

Policy Evaluation The policy evaluation step in the policy iteration involves optimizing coefficients to solve (HJB-PDE). Given a policy $\hat{a} = \left(\hat{\pi}, \hat{\lambda}_P, \hat{\lambda}_R, \hat{\lambda}_C\right)$, the following least-

squares approach is applied to minimize the norm of (HJB-PDE). This minimization is enabled by the linearity of the value function and its derivatives in $\mathbf{c}$. For brevity, we let $\hat{F}_k(\mathbf{c}^t)$ denote the vectorized value function estimate at all spatial states in $\boldsymbol{y}_P \times \boldsymbol{y}_R \times \boldsymbol{y}_I$ and let the $\partial_D \hat{F}_k(\mathbf{c})$ notation represent partial derivatives.

$$
\begin{aligned}
\underset{\mathbf{c}}{\text{minimize}} \quad & \Bigg\| \frac{\hat{F}_k^+ - \hat{F}_k(\mathbf{c})}{\Delta_t} - r\hat{F}_k(\mathbf{c}) + \tilde{u}_k \left((1-\nu_I)\hat{\lambda}_C(t) X_I(t) \right) \\
& + \partial_P \hat{F}_k(\mathbf{c}) \left((r(1 - \hat{\pi}^T(t)\mathbb{1}_N) + \hat{\pi}^T(t)\mu) X_P(t) + (1-\nu_I)\hat{\lambda}_P(t) X_I(t) \right) \\
& + \partial_R \hat{F}_k(\mathbf{c}) \left(\mu_R X_R(t) + (1+\kappa)\hat{\lambda}_R(t) X_I(t) \right) + \partial_I \hat{F}_k(\mathbf{c}) \left(\mu_I X_I(t) \right) \\
& + \frac{1}{2}\partial_{PP}\hat{F}_k(\mathbf{c}) X_P(t)^2 \hat{\pi}^T(t) C_{PP} \hat{\pi}(t) + \frac{1}{2}\partial_{RR}\hat{F}_k(\mathbf{c}) X_R(t)^2 C_{RR} + \frac{1}{2}\partial_{II}\hat{F}_k(\mathbf{c}) X_I(t)^2 C_{II} \\
& + \partial_{PR}\hat{F}_k(\mathbf{c}) X_P(t) X_R(t) \hat{\pi}^T(t) C_{PR} + \partial_{PI}\hat{F}_k(\mathbf{c}) X_P(t) X_I(t) \hat{\pi}^T(t) C_{PI} \\
& + \partial_{RI}\hat{F}_k(\mathbf{c}) X_R(t) X_I(t) C_{RI} \Bigg\|_2^2 \\
\text{s.t.} \quad & \hat{F}_k(\mathbf{c}) \geq 0 \\
& \partial_D \hat{F}_k(\mathbf{c}) \geq 0 \quad D \in \mathbf{D} \\
& \partial_{DD} \hat{F}_k(\mathbf{c}) \leq 0 \quad D \in \mathbf{D} \\
& \hat{F}_k(\mathbf{c}) \Big|_{X_P = X_R = X_I = 0} = 0
\end{aligned}
\tag{2.59}
$$

where $\hat{F}_k^+$ represents the vectorized value function estimate at time $t + \Delta_t$. Here, the first set of conditions ensures nonnegativity of the value function estimate. The second set of conditions ensures the value function estimate is nondecreasing. The third set of conditions enforces concavity of the value function estimate in each spatial dimension. The final condition ensures the value-to-go is zero when the investor has no portfolio value, no retirement value, and no income. We assume the constraints applied at the Chebyshev nodes are sufficient to ensure nonnegativity, monotonicity, and concavity throughout the entire domain. This problem has $(d_P+1)(d_R+1)(d_I+1)$ decision variables and $7(m_P m_R m_I)+1$ constraints.

Policy Improvement Given an estimate of the value function, denoted by $\hat{F}_k(\mathbf{c})$, the policy improvement step can be performed by maximizing H in (HJB-PDE). This is achieved through the following convex programming formulation, applied pointwise for all $(y_P, y_R, y_I) \in \boldsymbol{y}_P \times \boldsymbol{y}_R \times \boldsymbol{y}_I$:

$$\begin{aligned}
\underset{\pi,\,\lambda_P,\,\lambda_R}{\text{minimize}} \quad & -\tilde{u}_k\left((1-\nu_I)(1-\lambda_P-\lambda_R)y_I\right) - \partial_P \hat{F}_k(\mathbf{c})\left((\mu - r\mathbb{1}_N)^T \pi y_P + (1-\nu_I)\lambda_P y_I\right) \\
& - \partial_R \hat{F}_k(\mathbf{c})\left((1+\kappa)\lambda_R y_I\right) - \frac{1}{2}\partial_{PP} y_P^2 \pi^T C_{PP} \pi - \partial_{PR} y_P y_R \pi^T C_{PR} \\
& - \partial_{PI} \hat{F}_k(\mathbf{c}) y_P y_I \pi^T C_{PI} \\
\text{s.t.} \quad & \lambda_P + \lambda_R \leq 1, \\
& \pi^T \mathbb{1}_N \leq 1, \\
& \lambda_R \leq \gamma, \\
& \pi^T C_{PP} \pi \leq \sigma_{\max}^2, \\
& \pi, \lambda_P, \lambda_R \geq 0
\end{aligned} \tag{2.60}$$

2.3.3.2 Computational Complexity

We may express the computational complexity of our approach as

$$\frac{t_K}{\Delta_t}\left(n\left(\xi_{\text{eval}}(d_P, d_R, d_I, m_P, m_R, m_I) + m_P m_R m_I \xi_{\text{impr}}(N)\right)\right) + (K-1)\, m_P m_R m_I \xi_{\text{GC}} \tag{2.61}$$

where n is the average number of policy iterations, $\xi_{\text{eval}}(\cdot)$ is the complexity of solving (2.59), $\xi_{\text{impr}}(\cdot)$ is the complexity of solving (2.60), and ξ_{GC} is the complexity of optimizing the goal contribution at a single state.

In the following subsection, the parameters and utility functions used to solve the problem are presented.

2.3.4 Problem Parameterization

In this section, we present our numerical experiments. Our experiments begin with solving a "baseline" case for a typical investor, followed by some comparative studies to measure the sensitivity of the optimal decisions to the situations of both the investor and the market. All of our experiments use $m_P = m_R = m_I = 30$ and $d_P = d_R = d_I = 6$ for the Chebyshev polynomials.

2.3.4.1 The Baseline Investor

Our hypothetical client invests within the following context. Their income is taxed at a rate of 15%. Half of their net pay is spent on necessities, resulting in $\nu_I = 0.575$. We set a portfolio tax of 15%, aligned with the Internal Revenue Service (IRS) capital gain tax rates in 2023 [3]. The investor's income is assumed to grow at 2% per year on average, in line with historical inflation targeting [83]. We assume there are no taxes on retirement withdrawals.

The investor intends to retire in 40 years with $1,000,000 and has two goals before then: a $250,000 expense in ten years and a $500,000 expense in twenty years. The investor considers their second goal to be twice as important as the first goal, and retirement to be twice as important as the second goal.

The investor's retirement contribution is capped at 10%. For an investor with roughly the median income of full-time workers in the USA in 2022, this is roughly equivalent to the $6,000 limit for IRA contributions in 2022 at the initial time of the problem [1, 18]. A one-to-one match is made by the investor's employer, IBM, who provides a third of the investor's compensation in the form of company stock. The investor's chosen retirement fund is the Vanguard Target Retirement 2050 Fund ($VFIFX).

In addition to a bank account that returns 3% a year, the investor allocates their portfolio funds between stocks and bonds, represented by the SPDR S&P 500 ETF ($SPY) and the iShares 20 Plus Year Treasury Bond ETF ($TLT). A maximum portfolio volatility of 15% is chosen, with a rebalance period of one year.

We parameterize the model using the above description as follows. We estimate the asset drifts and covariances using their yearly log-returns of adjusted close prices from 2007 to 2022, inclusive:

$$\mu = \begin{bmatrix} 0.0361 \\ 0.0819 \end{bmatrix} \begin{matrix} \text{\$TLT} \\ \text{\$SPY} \end{matrix} \tag{2.62}$$

$$\mu_R = 0.0602 \tag{2.63}$$

$$C = \begin{matrix} \\ \text{\$TLT} \\ \text{\$SPY} \\ \text{\$VFIFX} \\ \text{I} \end{matrix} \begin{matrix} \begin{matrix} \text{\$TLT} & \text{\$SPY} & \text{\$VFIFX} & \text{I} \end{matrix} \\ \begin{bmatrix} 0.0335 & -0.0050 & -0.0053 & -0.0037 \\ -0.0050 & 0.0371 & 0.0335 & 0.0058 \\ -0.0053 & 0.0335 & 0.0312 & 0.0058 \\ -0.0037 & 0.0058 & 0.0058 & 0.0120 \end{bmatrix} \end{matrix} \tag{2.64}$$

2.3.4.2 Utility Functions

We assume the investor obtains linear utility for each goal contribution until their target contribution is reached, after which they obtain zero marginal utility. Utilities are weighted by their relative importances:

$$u_1(x) = \frac{1}{4} \min\left\{\frac{x}{250000}, 1\right\} \tag{2.65}$$

$$u_2(x) = \frac{1}{2} \min\left\{\frac{x}{500000}, 1\right\} \tag{2.66}$$

$$u_3(x) = \min\left\{\frac{x}{1000000}, 1\right\}. \tag{2.67}$$

In Section 2.3.3.1, Figure 2.4 visualizes the terminal value at retirement for this investor, having liquidated their assets and obtaining a utility of $u_3(0.85X_P(T) + X_R(T))$.We likewise assume that, for the entire horizon, the investor receives linear utility for consumption up to a consumption target of \$60,000 per year. Were the investor to satisfy this target throughout

their career, we assume that the total discounted utility from consumption would equal their discounted maximum retirement utility. This implies

$$\tilde{u}_k(x) = \frac{r}{e^{rT}-1} \min\left\{\frac{x}{60000}, 1\right\}, \quad k = 1, 2, ..., K. \tag{2.68}$$

Collectively, we denote the goal and consumption targets to be $(\varphi_1, \varphi_2, \varphi_3, \varphi_C) = (50000, 300000, 1000000, 60000)$, and the goal priorities to be $(\alpha_1, \alpha_2, \alpha_3, \alpha_C) = (1/4, 1/2, 1, r/(e^{rT}-1))$.

The value function has a maximum under this parameterization and thus (2.37) becomes

$$\max_X V_k(t, X) = \frac{r}{e^{rT}-1} \int_t^{t_K} e^{-rt_K} dt + \sum_{k=\tilde{k}(t)}^{K} e^{-rt_k} u_k(\varphi_k), \quad t \in [t_{k-1}, t_k), \quad k = 1, 2, ..., K \tag{2.69}$$

2.3.4.3 Safe Levels of Portfolio and Retirement Value

In prior work, the notion of a "safe level" of wealth is defined that guarantees all future utility [21]. This concept is helpful for contextualizing model behavior. Finite safe levels for full utility do not exist in our model because of the income process stochasticity. Instead, we establish an analogous concept of safety by comparing the maximum of the value function to the expected remaining utility if the investor were to disengage with the market as much as possible. For the remainder of their career, the investor would allocate their portfolio entirely to the risk-free asset and consume their entire income. We can determine the value-to-go under this strategy as follows. Let a_d be the disengaged investor's action (i.e. $\lambda_C^{a_d} = 1$, $\pi_B^{a_d} = 1$). Let $\tilde{G}_k \in \mathbb{R}_{\geq 0}$ denote the withdrawal from (exclusively) the portfolio towards goal $k \in \left\{\tilde{k}(t), \tilde{k}(t)+1, ..., K\right\}$. We then have the following convex set of feasible future portfolio

withdrawals under a_d as a function of t and $X_P(t)$:

$$\Gamma(t, X_P(t)) = \left\{ (\tilde{G}_{\tilde{k}(t)}, ..., \tilde{G}_K) > 0 \;\middle|\; X_P(t) - \sum_{k=\tilde{k}(t)}^{i} e^{-r(t_k - t)} \tilde{G}_k \geq 0, i = \tilde{k}(t), ..., K \right\} \quad (2.70)$$

where the condition ensures there is no overdraft on the portfolio. We then have the following theorem.

Theorem 4. *(Disengaged Investor's Value-to-go) For the parameterization described above, let $V^d(t, X)$ denote the value-to-go of the disengaged investor. We then have*

$$\begin{aligned} V^d(t, X) = \max_{\tilde{G}_{\tilde{k}(t)}, ..., \tilde{G}_K \in \Gamma(t, X_P)} & \left(\sum_{k=\tilde{k}(t)}^{K-1} e^{-r(t_k - t)} u_k((1 - \nu_P)\tilde{G}_k) - \frac{e^{-r(T-t)}}{\varphi_K} A_1(t, X_R, \tilde{G}_K) \right) \\ & + A_2(t, X_I) + e^{-r(T-t)} \end{aligned} \quad (2.71)$$

$$\begin{aligned} A_1(t, X_R, \tilde{G}_K) = & \left(\varphi_K - (1 - \nu_P)\tilde{G}_K \right) \Phi \left(\frac{\ln \frac{\varphi_K - (1-\nu_P)\tilde{G}_K}{X_R} - \left(\mu_R - \frac{C_{RR}}{2} \right)(T - t)}{\sqrt{C_{RR}(T - t)}} \right) \\ & - X_R e^{\mu_R (T-t)} \Phi \left(\frac{\ln \frac{\varphi_K - (1-\nu_P)\tilde{G}_K}{X_R} - \left(\mu_R + \frac{C_{RR}}{2} \right)(T - t)}{\sqrt{C_{RR}(T - t)}} \right) \end{aligned} \quad (2.72)$$

$$A_2(t, X_I) = \alpha_C \int_t^T e^{-r(s-t)} \left(1 - \frac{1}{\varphi_C} A_3(t, s, X_I) \right) ds \quad (2.73)$$

$$\begin{aligned} A_3(t, s, X_I) = & \varphi_C \Phi \left(\frac{\ln \left(\frac{\varphi_C}{(1-\nu_I)X_I} \right) - \left(\mu_I - \frac{C_{II}}{2} \right)(s - t)}{\sqrt{C_{II}(s - t)}} \right) \\ & - X_I e^{\mu_I (s-t)} \Phi \left(\frac{\ln \left(\frac{\varphi_C}{(1-\nu_I)X_I} \right) - \left(\mu_I + \frac{C_{II}}{2} \right)(s - t)}{\sqrt{C_{II}(s - t)}} \right) \end{aligned} \quad (2.74)$$

for $X_P \geq 0$, $X_R > 0$, and $X_I > 0$, where Φ is the cumulative distribution function of the standard normal distribution. We cover the edge cases for when $X_R = 0$ or $X_I = 0$ with: $A_1(t, 0, \tilde{G}_K) = \varphi_K - (1 - \nu_P)\tilde{G}_K \quad \forall t, \tilde{G}_K$ and $A_3(t, s, 0) = \varphi_C \quad \forall t$ and $s > t$.

Proof. We have, for all $t \in [0, T), X \geq 0$,

$$\begin{aligned} V^d(t,X) = \max_{\tilde{G}_{\tilde{k}(t)},\ldots,\tilde{G}_K \in \Gamma(t,X_P)} \Bigg(& \mathbb{E}^{a_d} \Bigg[\int_t^T e^{-r(s-t)} \tilde{u}((1-\nu_I) X_I(s))\, ds \\ & + \sum_{k=\tilde{k}(t)}^{K-1} e^{-r(t_k - t)} u_k((1-\nu_P)\tilde{G}_k) \\ & + e^{-r(T-t)} u_K((1-\nu_P)\tilde{G}_k + X_R(T)) \Bigg] \Bigg) \end{aligned} \tag{2.75}$$

$$\begin{aligned} = \mathbb{E}^{a_d} & \left[\int_t^T e^{-r(s-t)} \tilde{u}((1-\nu_I) X_I(s))\, ds \right] \\ & + \max_{\tilde{G}_{\tilde{k}(t)},\ldots,\tilde{G}_K \in \Gamma(t,X_P)} \Bigg(\sum_{k=\tilde{k}(t)}^{K-1} e^{-r(t_k - t)} u_k((1-\nu_P)\tilde{G}_k) \\ & + e^{-r(T-t)} \left(1 - \frac{1}{\varphi_K} \mathbb{E}^{a_d} \left[\max \left\{ 0, \varphi_K - (1-\nu_P)\tilde{G}_k - X_R(T) \right\} \right] \right) \Bigg) \end{aligned} \tag{2.76}$$

via linearity of expectation. We may use the maximum instead of supremum due to the compactness of Γ and the continuity of the goal utility functions and A_1. The expected utility from the retirement goal follows from

$$\begin{aligned} & \mathbb{E}^{a_d} \left[\max \left\{ 0, \varphi_K - (1-\nu_P)\tilde{G}_k - X_R(T) \right\} \right] \\ & = \int_0^{\varphi_K - (1-\nu_P)\tilde{G}_K} \varphi_K - (1-\nu_P)\tilde{G}_K - x\, dF_R(x) \end{aligned} \tag{2.77}$$

$$= \int_0^{\varphi_K - (1-\nu_P)\tilde{G}_K} \varphi_K - (1-\nu_P)\tilde{G}_K dF_R(x) - \int_0^{\varphi_K - (1-\nu_P)\tilde{G}_K} x dF_R(x) \tag{2.78}$$

$$\begin{aligned} & = \left(\varphi_K - (1-\nu_P)\tilde{G}_K \right) F_R \left(\varphi_K - (1-\nu_P)\tilde{G}_K \right) \\ & \quad + \mathbb{E}^{a_d} \left[X_R(T) \mid X_R(T) < \varphi_K - (1-\nu_P)\tilde{G}_k \right] \mathbb{P} \left(X_R(T) < \varphi_K - (1-\nu_P)\tilde{G}_K \right) \end{aligned} \tag{2.79}$$

where $F_R(x) = \mathbb{P}\left(X_R(T) \leq x \mid \mathcal{F}_t\right)$. As $\lambda_R^{a_d} = 0$ and the expectation of retirement utility is independent of X_P and X_I, (2.9) simplifies to

$$dX_R(t) = \mu_R X_R(t)\, dt + X_R(t) \sqrt{C_{RR}} d\mathbf{W}_R(t) \tag{2.80}$$

where $\mathbf{W}_R$ represents the individual Brownian motion of the retirement asset process. We therefore have that $X_R(T)$ is log-normally distributed:

$$\ln X_R(T) \sim \mathcal{N}\left(\ln X_R(t) + \left(\mu_R - \frac{1}{2}C_{RR}\right)(T-t), C_{RR}(T-t)\right). \tag{2.81}$$

It follows that

$$\begin{aligned}
&\mathbb{E}^{a_d}\left[X_R(T) \;\middle|\; X_R(T) < \varphi_K - (1-\nu_P)\tilde{G}_k\right]\mathbb{P}\left(X_R(T) < \varphi_K - (1-\nu_P)\tilde{G}_K\right) \\
&\quad = \begin{cases} X_R(t)e^{\mu_R(T-t)}\Phi\left(\frac{\ln\left(\frac{\varphi_K-(1-\nu_P)\tilde{G}_k}{X_R(t)}\right)-\left(\mu_R+\frac{C_{RR}}{2}\right)(T-t)}{\sqrt{C_{RR}(T-t)}}\right) & X_R(t) > 0 \\ 0 & X_R(t) = 0. \end{cases}
\end{aligned} \tag{2.82}$$

We also have

$$\begin{aligned}
&\left(\varphi_K - (1-\nu_P)\tilde{G}_K\right)F_R\left(\varphi_K - (1-\nu_P)\tilde{G}_K\right) \\
&\quad = \begin{cases} \left(\varphi_K - (1-\nu_P)\tilde{G}_K\right)\Phi\left(\frac{\ln\left(\frac{\varphi_K-(1-\nu_P)\tilde{G}_K}{X_R(t)}\right)-\left(\mu_R-\frac{C_{RR}}{2}\right)(T-t)}{\sqrt{C_{RR}(T-t)}}\right) & X_R(t) > 0 \\ \varphi_K - (1-\nu_P)\tilde{G}_K & X_R(t) = 0. \end{cases}
\end{aligned} \tag{2.83}$$

For the consumption utility, we invoke the stochastic Fubini theorem to exchange the order of integration. This requires the consumption utility function meets sufficient measurability conditions [90]. Let $\psi : X_I \times [t,T] \times \Omega_I \to \mathbb{R}$, in which Ω_I is the sample space of paths of X_I, represent consumption utility. We then have

$$\int_0^\infty \left(\int_t^T |\psi(i,s)|^2\, ds\right)^{\frac{1}{2}} dF_I(i) \leq \int_0^\infty \left(\int_t^T \alpha_C^2\, ds\right)^{\frac{1}{2}} dF_I(i) \tag{2.84}$$

$$= \alpha_C\sqrt{T-t} \tag{2.85}$$

$$< \infty \tag{2.86}$$

where F_I is the distribution of paths of I. The finiteness of this integral then ensures

$$\int_0^\infty \int_t^T \psi(i,s)\, dX_I(s)\, dF_I(i) = \int_t^T \int_0^\infty \psi(i,s)\, dF_I(i)\, dX_I(s) \tag{2.87}$$

which implies

$$\mathbb{E}^{a_d}\left[\int_t^T e^{-r(s-t)}\tilde{u}((1-\nu_I)X_I(s))\, ds\right] \tag{2.88}$$

$$= \int_t^T e^{-r(s-t)}\mathbb{E}^{a_d}\left[\tilde{u}((1-\nu_I)X_I(s))\right] ds \tag{2.89}$$

$$= \alpha_C \int_t^T e^{-r(s-t)}\left(1 - \frac{1}{\varphi_C}\mathbb{E}^{a_d}\left[\max\left\{0, \varphi_C - (1-\nu_I)X_I(s)\right\}\right]\right) ds. \tag{2.90}$$

Furthermore,

$$\mathbb{E}^{a_d}\left[\max\left\{0, \varphi_C - (1-\nu_I)X_I(s)\right\}\right] \tag{2.91}$$

$$= \int_0^{\frac{\varphi_C}{1-\nu_I}} \varphi_C - (1-\nu_I)x\, dF_I(x) \tag{2.92}$$

$$= \int_0^{\frac{\varphi_C}{1-\nu_I}} \varphi_C\, dF_I(x) - \int_0^{\frac{\varphi_C}{1-\nu_I}} (1-\nu_I)x\, dF_I(x) \tag{2.93}$$

$$= \varphi_C F_I\left(\frac{\varphi_C}{1-\nu_I}\right) - (1-\nu_I)\mathbb{E}^{a_d}\left[X_I(s) \;\middle|\; X_I(s) < \frac{\varphi_C}{1-\nu_I}\right]\mathbb{P}\left(X_I(s) < \frac{\varphi_C}{1-\nu_I}\right) \tag{2.94}$$

for $X_I(t) > 0$. Trivially, $A_2(t,0) = 0\ \forall t$. We have independence of X_I from X_P and X_R, and may therefore express (2.5) as

$$dX_I(t) = \mu_I X_I(t)\, dt + \sqrt{C_{II}} X_I(t) d\mathbf{W}_I(t) \tag{2.95}$$

where $\mathbf{W}_I$ represents the individual Brownian motion of the income process. Like the re-

tirement process, the income is therefore log-normally distributed:

$$\ln X_I(s) \sim \mathcal{N}\left(\ln X_I(s) + \left(\mu_I - \frac{1}{2}C_{II}\right)(s-t), C_{II}(s-t)\right), \quad s > t. \tag{2.96}$$

Therefore,

$$\begin{aligned}
&\mathbb{E}^{a_d}\left[X_I(s) \;\middle|\; X_I(s) < \frac{\varphi_C}{1-\nu_I}\right]\mathbb{P}\left(X_I(s) < \frac{\varphi_C}{1-\nu_I}\right) \\
&\quad = \begin{cases} X_I(t)e^{\mu_I(s-t)}\Phi\left(\frac{\ln\left(\frac{\varphi_C}{(1-\nu_I)X_I(t)}\right) - \left(\mu_I + \frac{C_{II}}{2}\right)(s-t)}{\sqrt{C_{II}(s-t)}}\right) & X_I(t) > 0 \\ 0 & X_I(t) = 0 \end{cases}
\end{aligned} \tag{2.97}$$

for $s > t$. We also have

$$\varphi_C F_I\left(\frac{\varphi_C}{1-\nu_I}\right) = \begin{cases} \varphi_C\Phi\left(\frac{\ln\left(\frac{\varphi_C}{(1-\nu_I)X_I(t)}\right) - \left(\mu_I - \frac{C_{II}}{2}\right)(s-t)}{\sqrt{C_{II}(s-t)}}\right) & X_I(t) > 0 \\ \varphi_C & X_I(t) = 0. \end{cases} \tag{2.98}$$

□

The argument of the maximum in (2.71) is concave in $\tilde{G}_k$, $k = \tilde{k}(t), \tilde{k}(t)+1, ..., K$. This because it is a non-negative weighted sum of concave functions, which may be seen from (2.76). The maximum argument can therefore be solved via the following convex optimization formulation:

$$\begin{aligned}
\underset{\tilde{G}_k,\, k=\tilde{k}(t),...,K}{\text{minimize}} \quad & -\left(\sum_{k=\tilde{k}(t)}^{K-1} e^{-r(t_k-t)}u_k((1-\nu_P)\tilde{G}_k) - \frac{e^{-r(T-t)}}{\varphi_K}A_1(t, X_R, \tilde{G}_K)\right) \\
\text{s.t.} \quad & X_P - \sum_{k=\tilde{k}(t)}^{i} e^{-r(t_k-t)}\tilde{G}_k \geq 0, \quad i = \tilde{k}(t), ..., K \\
& \tilde{G}_k \geq 0, \quad k = \tilde{k}(t), ..., K.
\end{aligned} \tag{2.99}$$

With V^d, we define levels of safety by comparing the disengaged investor's value-to-go to

the maximum of the value function over the spatial domain. The α-safe states for $t \in [0, T)$ are defined as

$$\chi_\alpha(t) = \left\{ X \in \mathbb{R}^3_{\geq 0} : \frac{V^d(t, X)}{\max_{X'} V(t, X')} \geq 1 - \alpha \right\}, \quad \alpha \in [0, 1]. \tag{2.100}$$

This measures the proportion of possible discounted future gain expected under the disengaged strategy, and is most similar to the concept of safety in [21]. It is easy to see that $V^d(t, x)$ is convex in x via a similar argument to Theorem 1. We then have that $\chi_\alpha(t)$ is a convex set for any $\alpha \in [0, 1]$ as, for $\beta^t_\alpha \triangleq (1 - \alpha) \sup_X V(t, X)$, we have that $\chi_\alpha(t)$ is the β^t_α-superlevel set of $V^d(t, X)$ [14]. As $V^d(t, X)$ is also nondecreasing in all components of X, therefore, for any $X^{(1)} \geq 0$, we have

$$X^{(1)} \in \chi_\alpha(t) \implies X^{(2)} \in \chi_\alpha(t), \quad \forall X^{(2)} \geq X^{(1)}, t \in [0, T), \alpha \in [0, 1] \tag{2.101}$$

where "$\geq$" refers to the component-wise partial ordering. These characteristics reduce the problem of identifying $\chi_\alpha(t)$ to finding its boundary, enabling faster calculation compared to evaluating the safety at each node.

2.3.4.4 Comparative Studies

We repeat the experiment with different parameterizations to compare an investor's decisions under different circumstances and demonstrate the flexibility of the model.

Increased Income Growth Rate We increase μ_I to 0.04 to infer the change in consumption behavior and risk-taking prior to goal realization.

No Employer Retirement Contribution We decrease κ to 0 to compare how decisions vary if an employer does not contribute anything to the employee's retirement. This experiment may also apply to self-employed investors.

High correlation between market and income We modify C to $\bar{C}$, described below, to increase the employee's income dependence on the market. That is, the correlation between X_I and the \$SPY ETF is increased. As the determinant of the correlation matrix must be positive, this yields $0.129 \leq \text{corr}(\$\text{SPY}, X_I) \leq 0.456$. For this experiment, we set $\text{corr}(\$\text{SPY}, X_I)$ to 0.45.

The following section presents the numerical results from the above experiments and a discussion on the optimal decisions of the investor.

2.4 Results and Discussion

We begin by presenting the results from the baseline investor, interpreting the model's recommendations at, and around, salient decisions. Following this, we examine the comparative studies to illuminate the sensitivity of the solutions to the model parameterization.

2.4.1 The Baseline Investor

2.4.1.1 Portfolio Near Retirement

Figure 2.5 below shows the allocations and volatility of the investor's portfolio one year prior to retirement for a trio of income values. For clarity, we exclude the 0.01-safe states from this figure and relevant following figures. In this figure, the level at which the investor fully satisfies their retirement goals is roughly the line from the top left corner of each heatmap to the bottom right as the investor's retirement target is met at a pre-tax portfolio withdrawal of \$1,180,000.

We see that the investor maintains a maximally volatile portfolio, consisting entirely of \$SPY and \$TLT, at lower levels of portfolio and retirement portfolio values. However, as their portfolio value and/or retirement portfolio values increase, the investor begins reallocating from both \$SPY and \$TLT to the risk-free asset. This is intuitive: the investor wishes to avoid potential losses, given the potential upside is limited.

The behavior changes as the investor's income increases. For combinations of high X_P and low X_R, the investor tends to choose a lower-volatility portfolio. This can be viewed as pricing in the coming year's income: the investor can afford to take less risk in the portfolio when more of the current shortfall can be effectively guaranteed by the income. Furthermore, the range of X_P and X_R values at which the investor should have less risky portfolio allocations is reduced for higher incomes, as evidenced by the low-X_P, high-X_R zone. For a lower income in this region, the investor should reduce portfolio volatility slightly to reduce potential losses. A higher-income investor, however, needn't make such concessions as full fundedness of their retirement goal is virtually guaranteed.

Figure 2.6 below shows how portfolio allocations change approaching retirement for investors with zero, moderate, and high incomes. Nearing retirement, for most X_P and X_R combinations close to the goal target, the portfolio volatility decreases. With additional time to retirement, the benefits of avoiding reduced utility are outweighed by the increased expected utility growth under a riskier portfolio.

For states with high X_P and very low X_R, the portfolio allocations become more risk-on approaching retirement. When relying almost entirely on the portfolio for retirement utility, it is sensible to trade off expected growth for portfolio safety given the increased time buffer over which portfolio gains may be made. We see this effect is more pronounced for higher incomes as, again, less risk needs to be taken when the investor's high income will likely carry their portfolio to the target within a few years.

2.4.1.2 Income Allocation Ten Years Until Retirement

Figure 2.7 below shows the income allocation ten years prior to retirement for a range of retirement account values. Having passed the second goal, the investor's savings are exclusively towards the retirement target. Examining the second row, we see that the investor will contribute the maximum allowable amount towards the retirement portfolio in most cases, aside for when either their retirement portfolio is well-funded (Figure 2.7(f)) or their port-

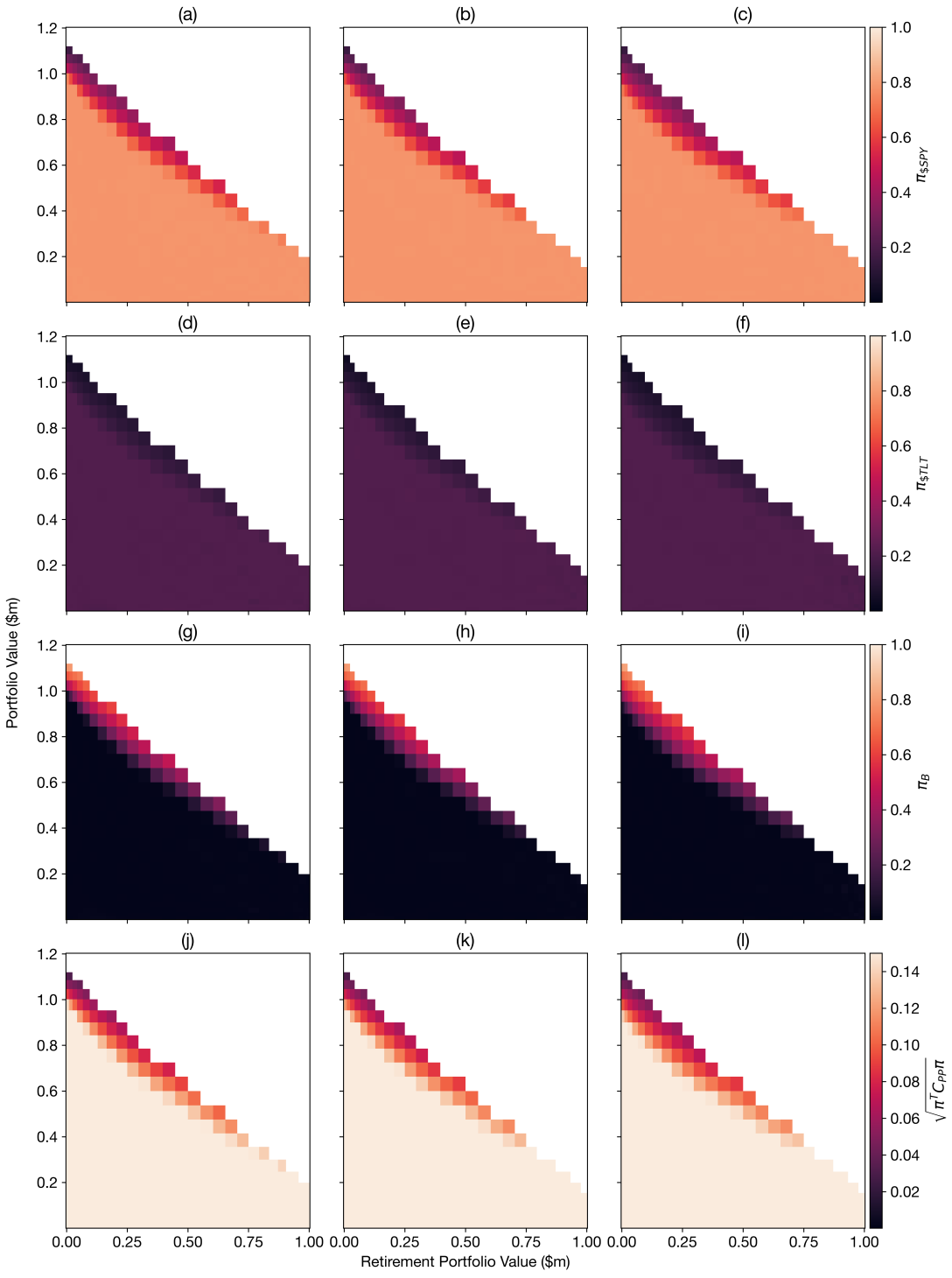

Figure 2.5: Baseline Investor Portfolio Volatility One Year Before Retirement per Income Level

Rows: $\pi_{\$\text{SPY}}$; $\pi_{\$\text{TLT}}$; π_B; $\sqrt{\pi^T C_{\text{PP}} \pi}$

Columns: $X_I \approx \$55{,}000$; $X_I \approx \$110{,}000$; $X_I \approx \$200{,}000$

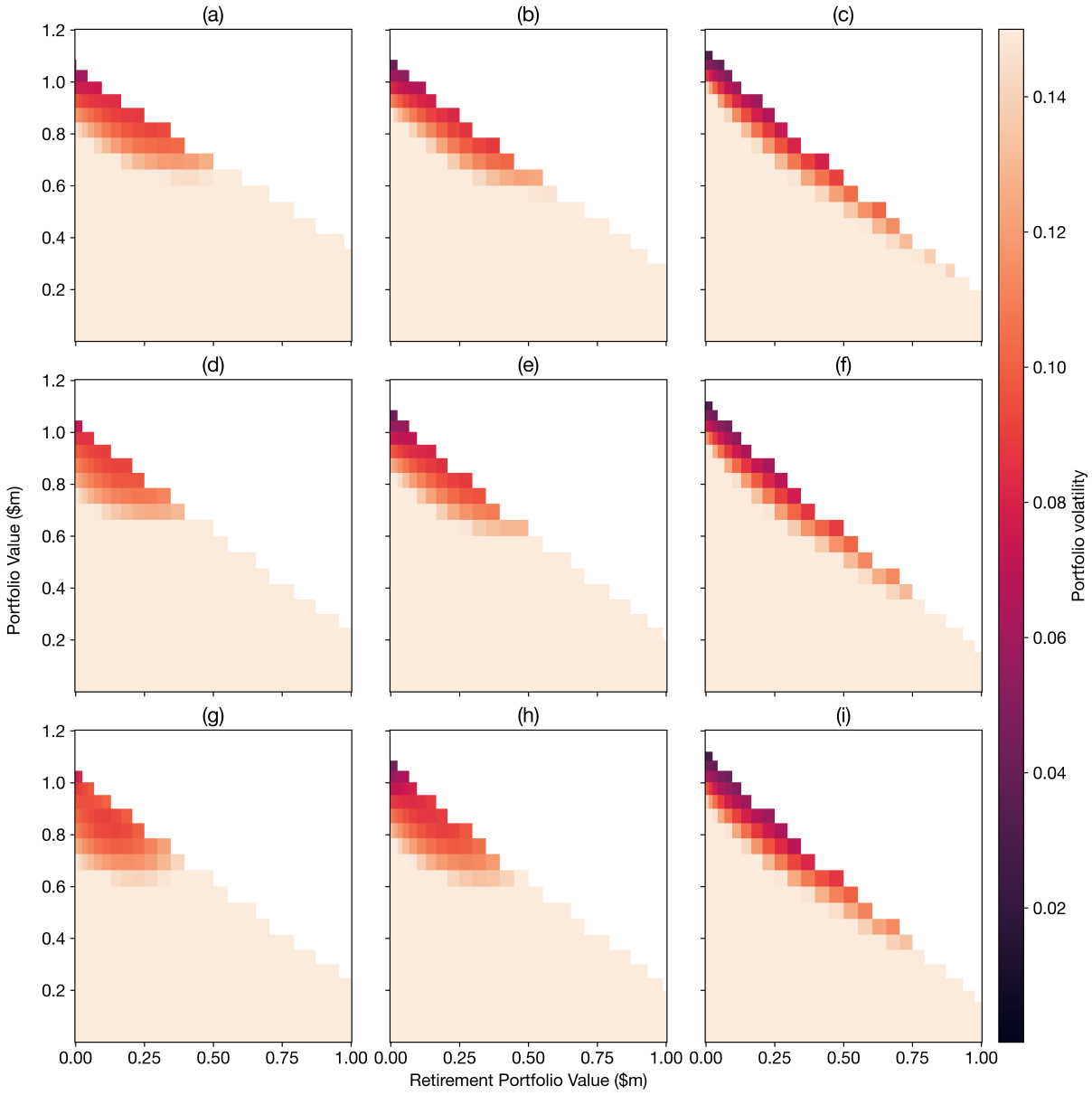

Figure 2.6: Baseline Investor Portfolio Volatility Approaching Retirement
Rows: $X_I = 0$; $X_I \approx \$55{,}000$; $X_I \approx \$200{,}000$
Columns: 3 years to retirement; 2 years to retirement; 1 year to retirement

folio value is so low compared with their retirement portfolio value that there are potential diversification benefits of contributing maximally to the portfolio (bottom of Figure 2.7(e)). In the former case, the investor opts to split their income entirely between the portfolio and immediate consumption, aside from the high-X_P, high-X_I region in which all future utility is effectively guaranteed and the λ_P-λ_R tradeoff is immaterial given that φ_C is consumed.

We also see a barrier partitioning the X_P-X_I plane into areas of consumption and saving. As both the investor's income and their retirement account value increase, they require less in their portfolio before they begin discretionary spending. This matches common advice to spend within one's means.

2.4.1.3 Income Allocation Leading up to the Second Goal

Figure 2.8 below shows the optimal income allocation leading up to, and immediately after, the second goal for an investor with an income of approximately $160,000 per year. The pre-tax contribution required to fully satisfy the goal is $588,235. Leading up to the goal deadline, if the investor does not have sufficient portfolio value to completely fund their second goal, they will contribute the large majority of their income towards the portfolio. With sufficient portfolio value, however, they will instead consume φ_C to maximize consumption utility, with the remainder divided between the portfolio and retirement.

Interestingly, when the investor's retirement portfolio is not well-funded, the investor opts to put a proportion of their income towards their retirement account even if their portfolio is insufficient to satisfy their second goal. The relative importance of retirement compared to this second goal is therefore demonstrated by this tradeoff. We also see a significant deviation in strategy following the goal contribution: the investor no longer places their income into their portfolio unless it is not well-funded, opting to consume a significant proportion immediately. With twenty years remaining until retirement, the investor can still comfortably consume without fear of lost terminal utility.

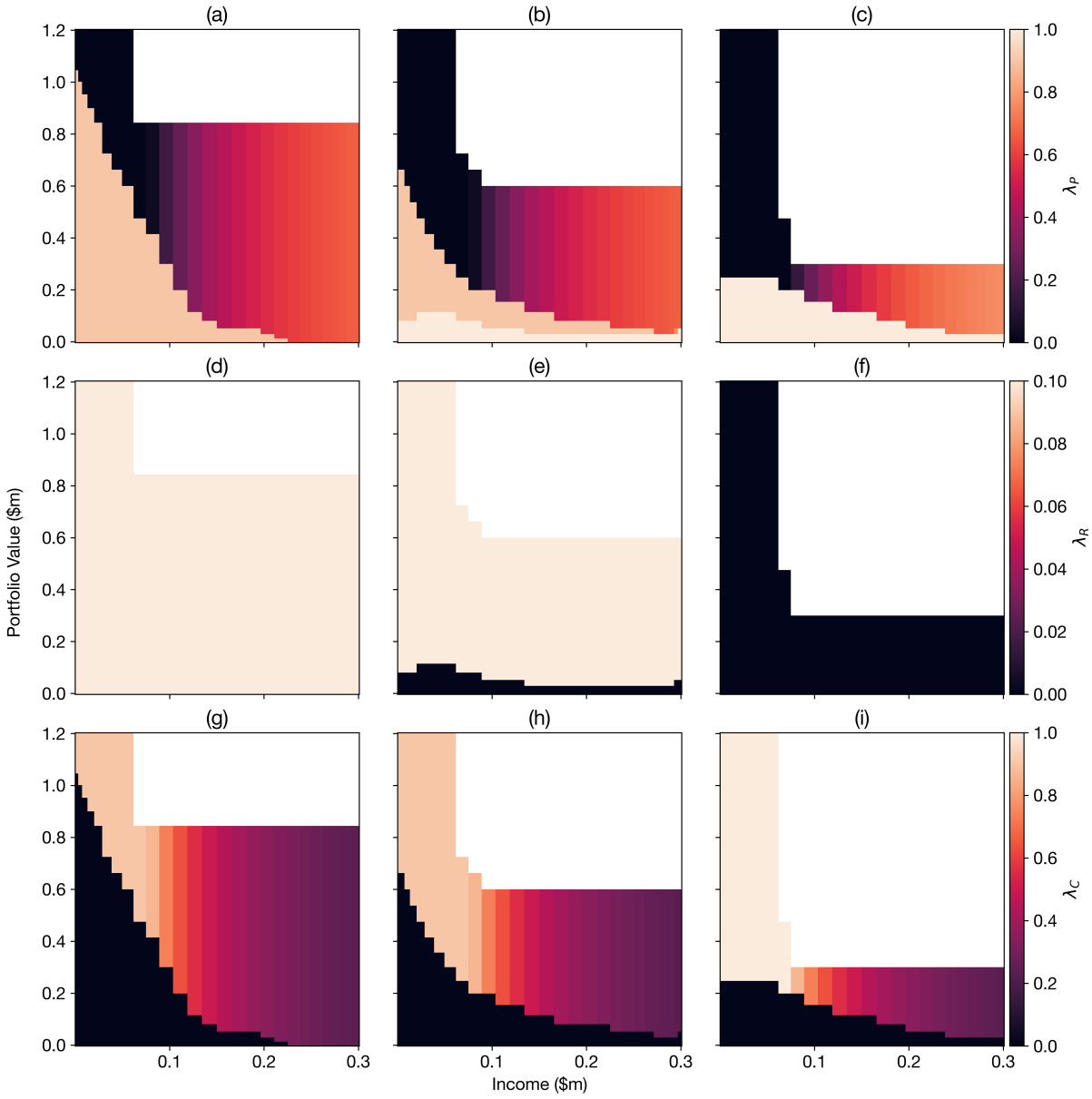

Figure 2.7: Baseline Investor Income Allocation 10 Years Before Retirement
Rows: λ_P; λ_R; λ_C
Columns: $X_R = 0$; $X_R \approx \$320{,}000$; $X_R \approx \$950{,}000$

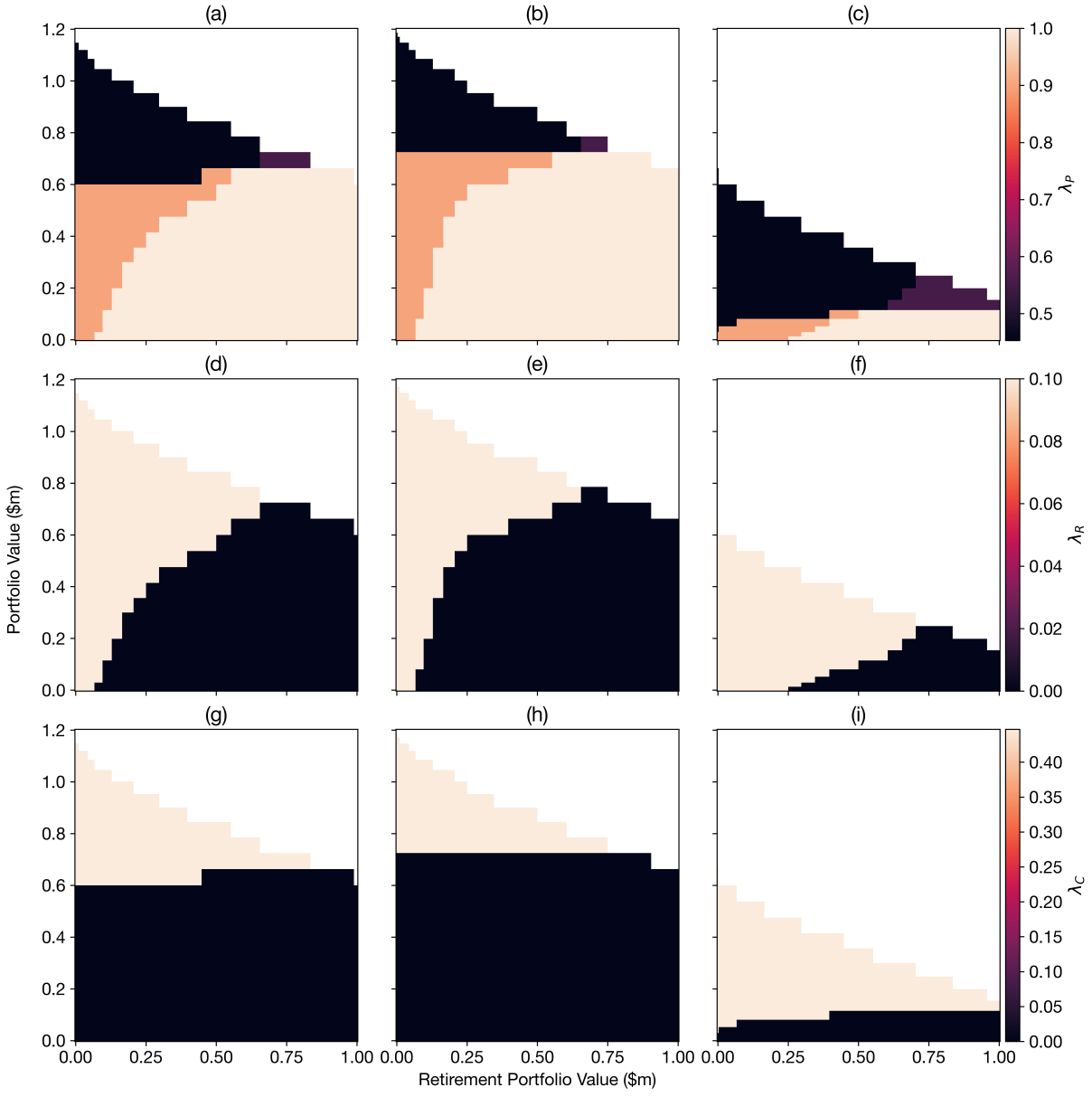

Figure 2.8: Baseline Investor Portfolio Contributions Before and After Goal 2 ($X_I \approx$ \$160,000)
Rows: λ_P; λ_R; λ_C
Columns: 2 years before goal 2; 1 year before goal 2; immediately after goal 2

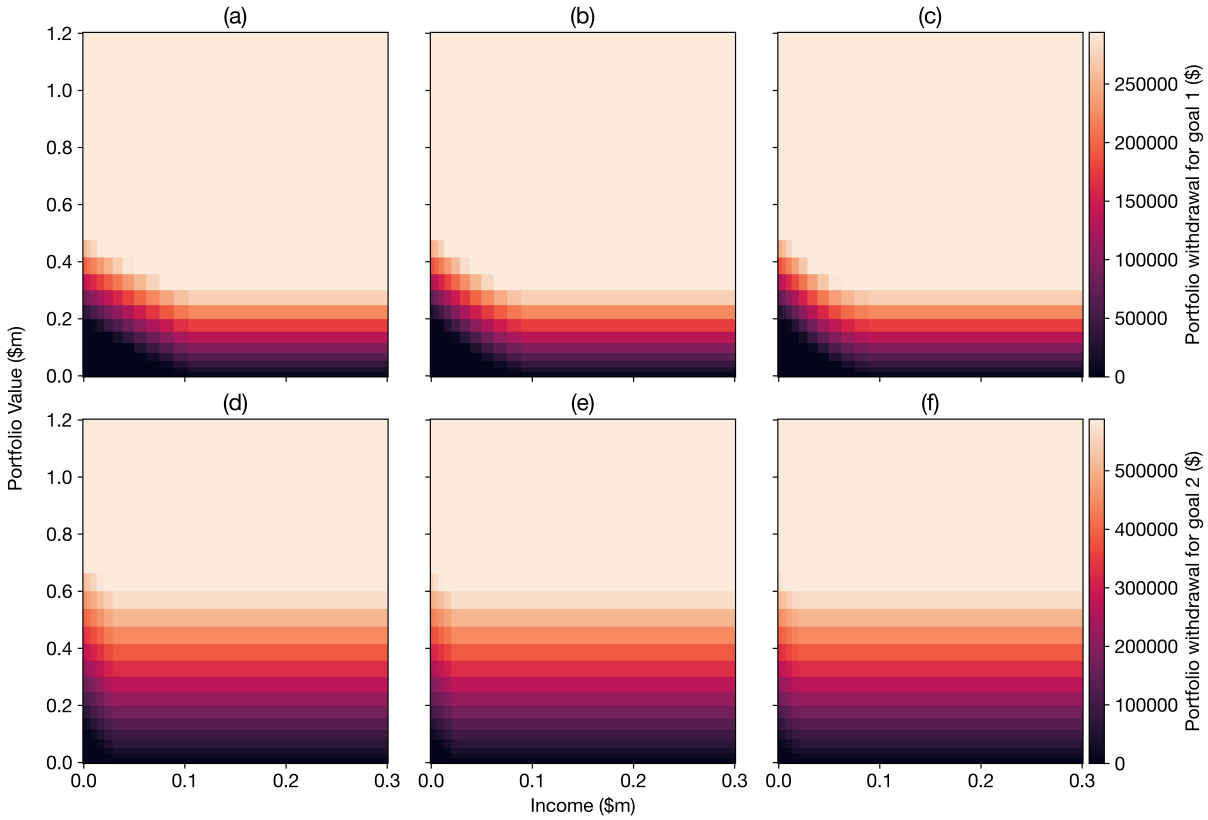

Figure 2.9: Baseline Investor Goal Contributions
Rows: $\bar{G}_1$; $\bar{G}_2$
Columns: $X_R \approx \$180{,}000$; $X_R \approx \$580{,}000$; $X_R = \$1{,}000{,}000$

2.4.1.4 Goal Contributions

Figure 2.9 below shows the investor's contributions towards each goal for a panel of retirement portfolio values. The investor generally withdraws enough to satisfy each goal's target, aside from when either: their portfolio value is too low, opting instead to withdraw their entire portfolio; or their income is low. For the latter case, the investor sometimes chooses to withhold a proportion of the goal target to save for future goal(s) even if they can reach the current goal's target.

2.4.2 Increased Market-Income Correlation

Intuitively, increasing the market-income correlation should lead to more risk-averse behavior. Poor market performance is now more likely to accompany a decrease in income, so the investor should expect a higher variance in outcomes. The following subsections confirm this intuition and demonstrate the impact on the investor's decision-making.

2.4.2.1 Change in Portfolio Near Retirement

Figure 2.10 below shows how the investor's optimal portfolio changes with the increased market-income correlation one year prior to retirement. For low-X_P, high-X_R states, as well as states generally close to the safe states, the investor with the higher market-income correlation has a lower optimal portfolio volatility. The increased likelihood of the investor's income and the S&P 500 moving in the same direction results in limited upside utility for a favorable outcome, but increased downside for an adverse outcome. This justifies a risk reduction.

However, some high-X_P, low-X_R states have an increased allocation to \$SPY. With a less correlated income to the market, the investor could reduce exposure to the market in expectation that their income could satisfy much of the utility shortfall during a market downturn. This is not the case with the increased correlation, as their income provides less of a hedge against lost utility. The investor should therefore remain more risk-on.

2.4.2.2 Change in Goal Contributions

Figures 2.11 and 2.12 below show the optimal goal contributions and how they differ from the baseline investor, respectively. For the second goal, we see a decrease in contributions for states with low X_I and moderate-to-low X_P, regardless of retirement portfolio value. This may be seen as avoiding retirement risk given the investor's situation is now less stable. At first glance, it may seem counterintuitive that the investor then contributes more to their first goal for some lower-X_I, low-X_P states. In this case, one can see the value of immediate

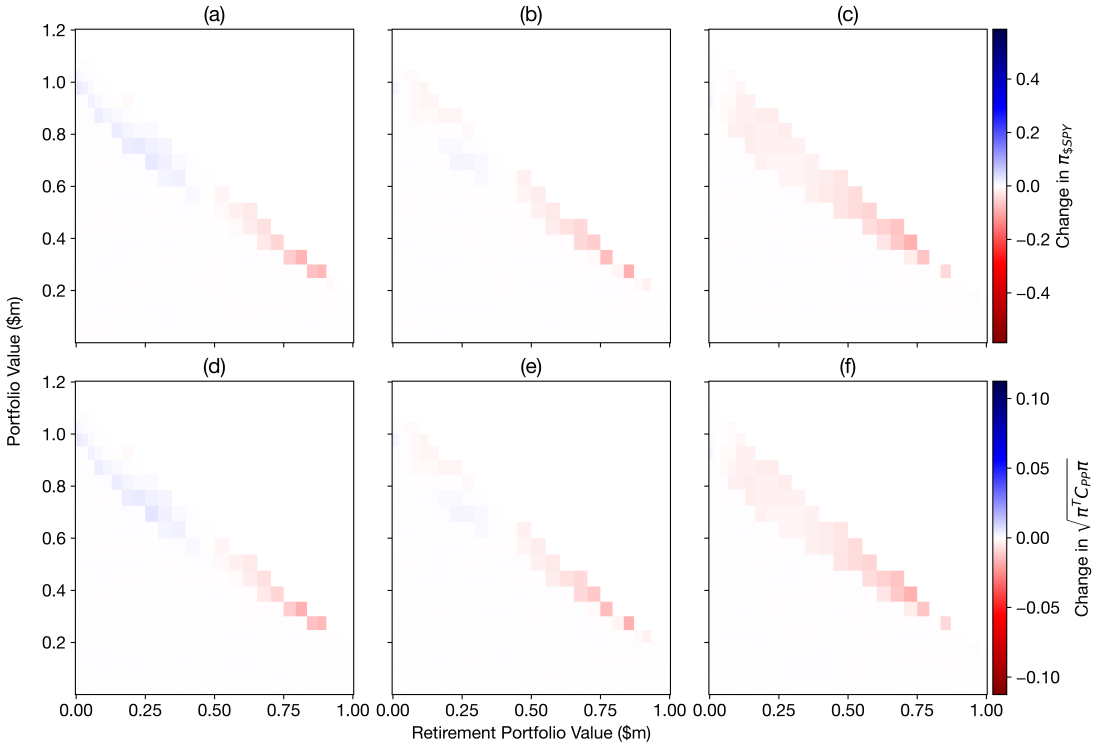

Figure 2.10: Change in Optimal Portfolio per Income Level when \$SPY-Income Correlation Increased, One Year Before Retirement
Rows: Change in $\pi_{\$\mathrm{SPY}}$; Change in $\sqrt{\pi^T C_{\mathrm{PP}} \pi}$
Columns: $X_I \approx \$55{,}000$; $X_I \approx \$110{,}000$; $X_I \approx \$200{,}000$

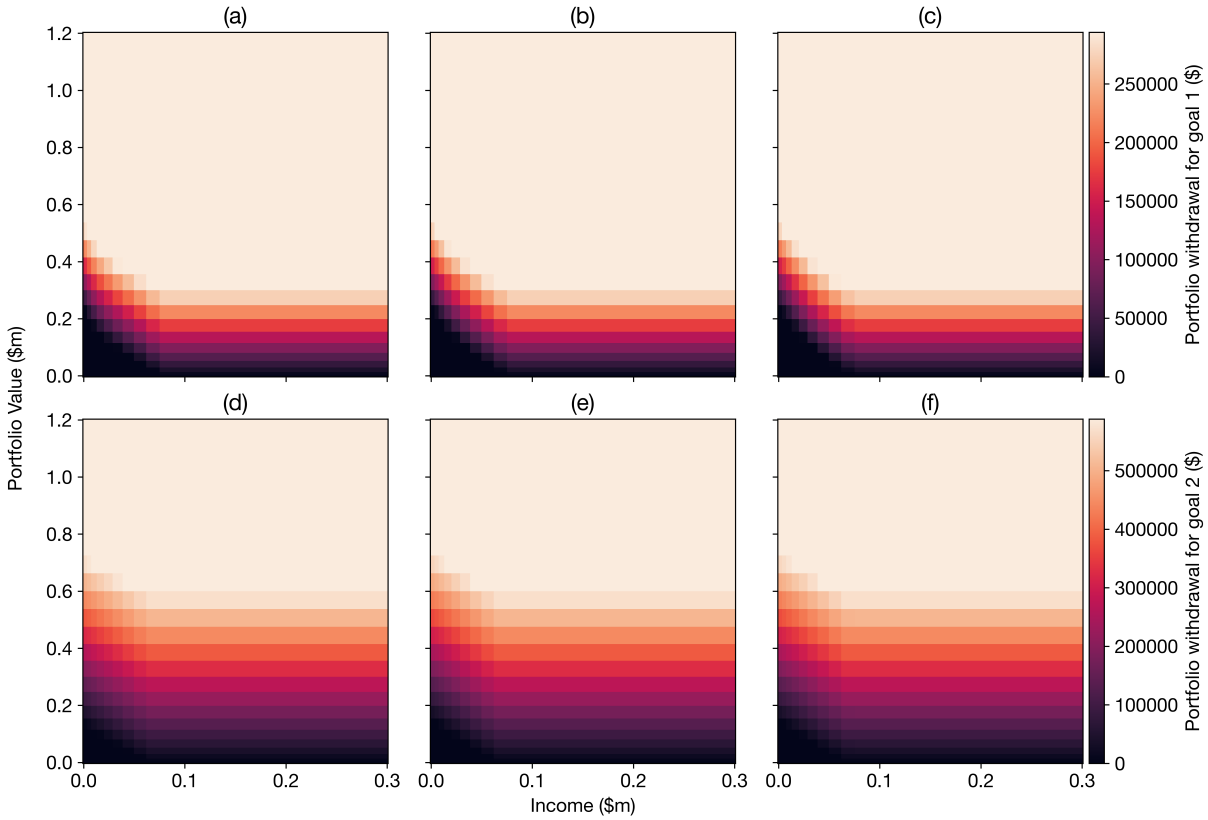

Figure 2.11: Increased Market-Income Correlation Investor Goal Contributions
Rows: $\bar{G}_1$; $\bar{G}_2$
Columns: $X_R \approx \$180{,}000$; $X_R \approx \$580{,}000$; $X_R = \$1{,}000{,}000$

utility when faced with increased future uncertainty.

2.4.2.3 Change in Income Allocation Leading up to the Second Goal

Figure 2.13 below shows the income allocation leading up to, and directly after, goal 2. Compared with Figure 2.8, the investor takes a more risk-averse approach, with limited consumption, bolstering the portfolio value and focusing more on retirement. For instance, the investor must now have more in their retirement account before ceasing retirement contribution. The additional impulse to contribute to retirement serves to hedge against adverse moves in the portfolio and income. The investor also reduces consumption to support the portfolio value both before and after the second goal time, even at higher portfolio values.

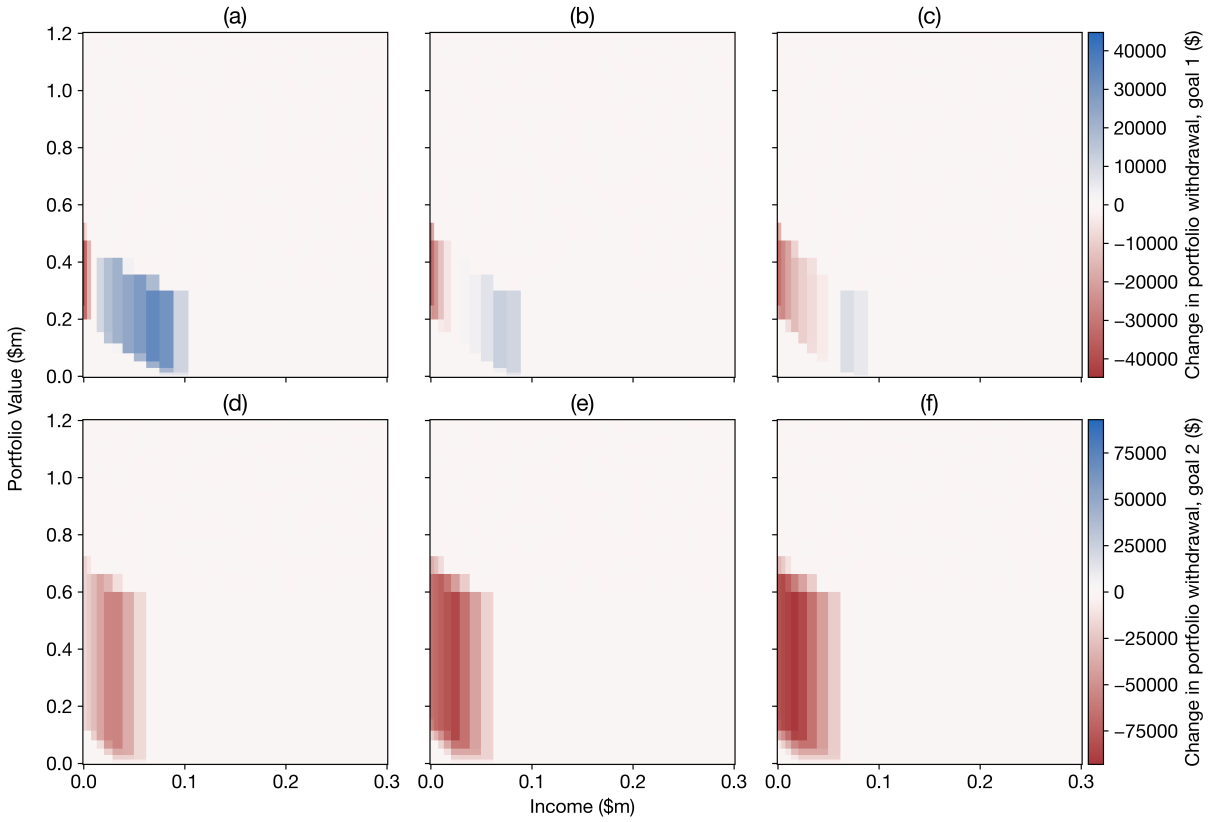

Figure 2.12: Difference Between Goal Contributions: Increased Market-Income Correlation Minus Baseline
Rows: Change in $\bar{G}_1$; Change in $\bar{G}_2$
Columns: $X_R \approx \$180{,}000$; $X_R \approx \$580{,}000$; $X_R = \$1{,}000{,}000$

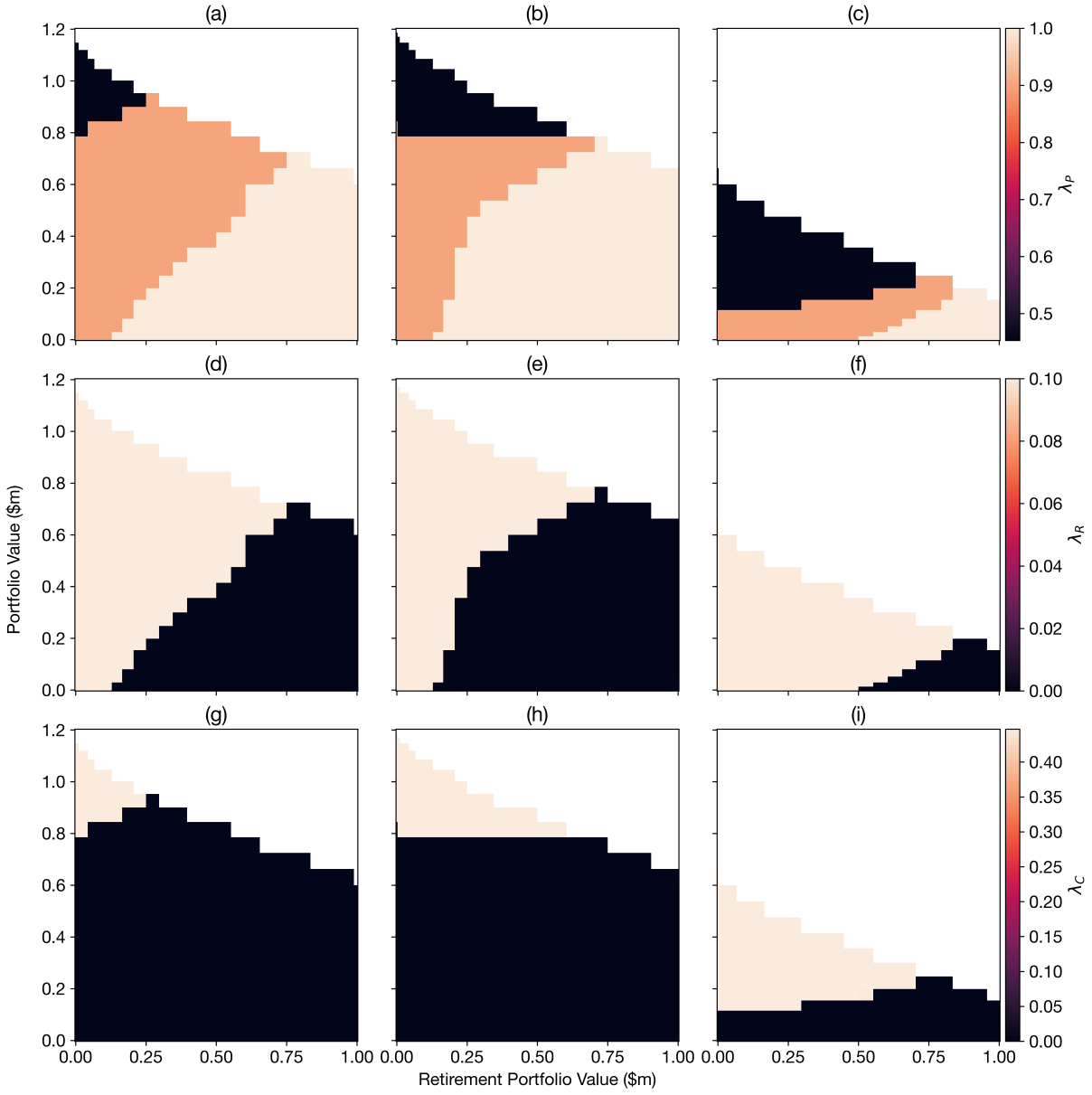

Figure 2.13: Increased Market-Income Correlation Investor Portfolio Contributions Before and After Goal 2 ($X_I \approx \$160{,}000$)
Rows: λ_P; λ_R; λ_C
Columns: 2 years before goal 2; 1 year before goal 2; immediately after goal 2

2.4.3 Zero Employer Retirement Contribution

The absence of employer contribution to the retirement account should disincentivize contributing to retirement and reduce the investor's expected terminal utility. The quantitative implications are shown in the following subsections.

2.4.3.1 Income Allocation Ten Years Until Retirement

Figure 2.14 below shows the investor's income allocation ten years prior to retirement. Compared with Figure 2.7(e), the investor now requires a generally higher portfolio value before consuming. Consumption is even curtailed at high-X_R values, demonstrated by the third column of plots. Evidently, the decreased drift in the retirement process threatens retirement utility enough to curtail consumption and encourage additional portfolio contributions.

Figure 2.14(e) reveals that, for a moderate retirement portfolio value, there are two X_P and X_I themes in which retirement contributions are recommended. The first, for moderate X_P and low X_I, simultaneously indicates that (1) the present value of marginal retirement contributions outweigh the immediate benefit of pure consumption, and (2) the diversification and/or tax benefits of marginal retirement contribution complement the portfolio sufficiently. In the other context, namely states with high X_I and satisfactory X_P, the investor consumes up to φ_C and hedges their portfolio with a maximal retirement contribution.

2.4.3.2 Change in Portfolio Near Retirement

Figure 2.15 below shows the change in portfolio allocation to the riskier asset and portfolio volatility compared with the baseline investor. Near safe levels, the investor chooses a comparatively increased portfolio volatility, a justifiable decision based on the reduced inflow from retirement contributions.

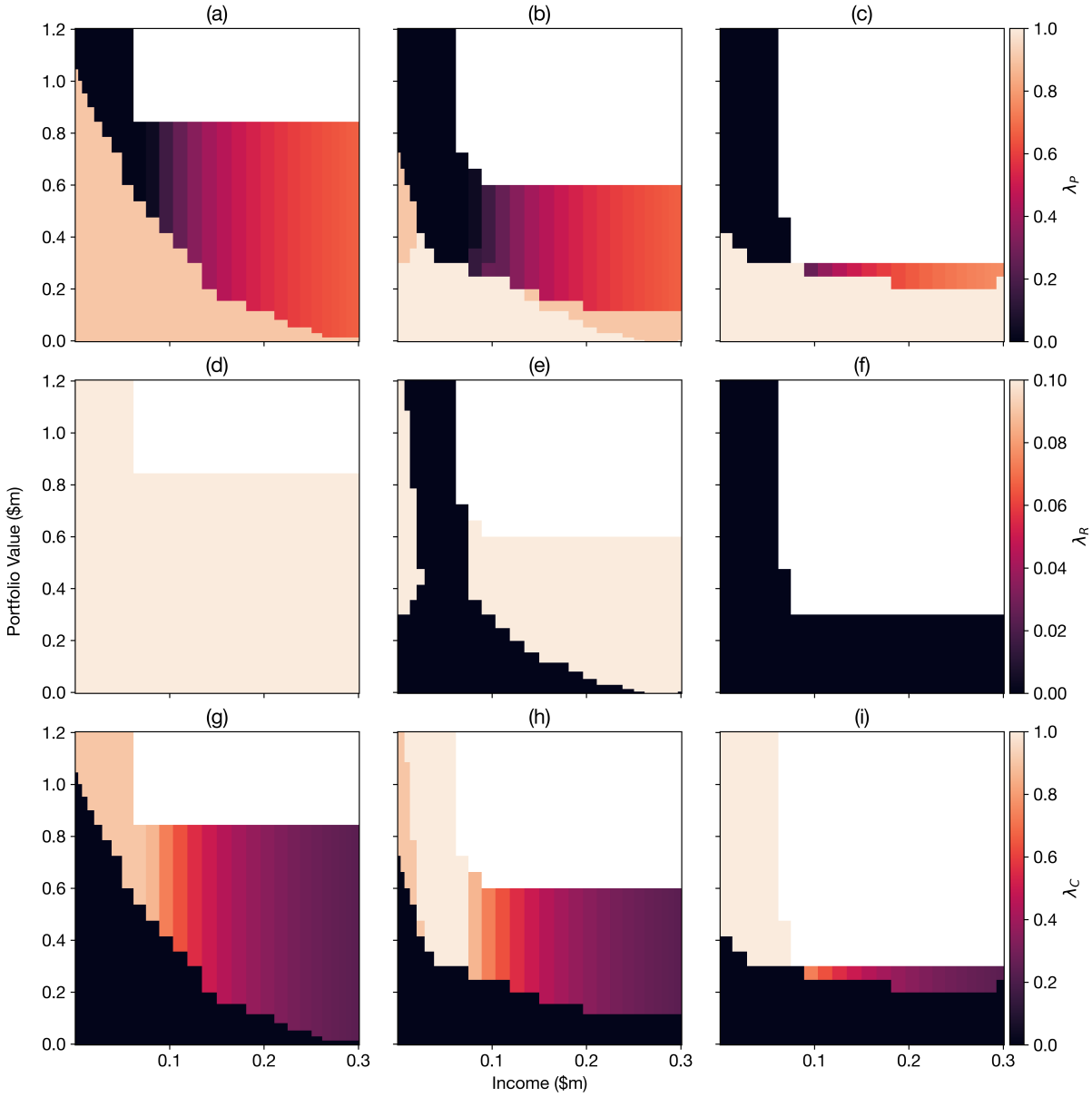

Figure 2.14: Zero Employer Contribution Investor Income Allocation 10 Years Before Retirement
Rows: λ_P; λ_R; λ_C
Columns: $X_R = 0$; $X_R \approx \$320{,}000$; $X_R \approx \$950{,}000$

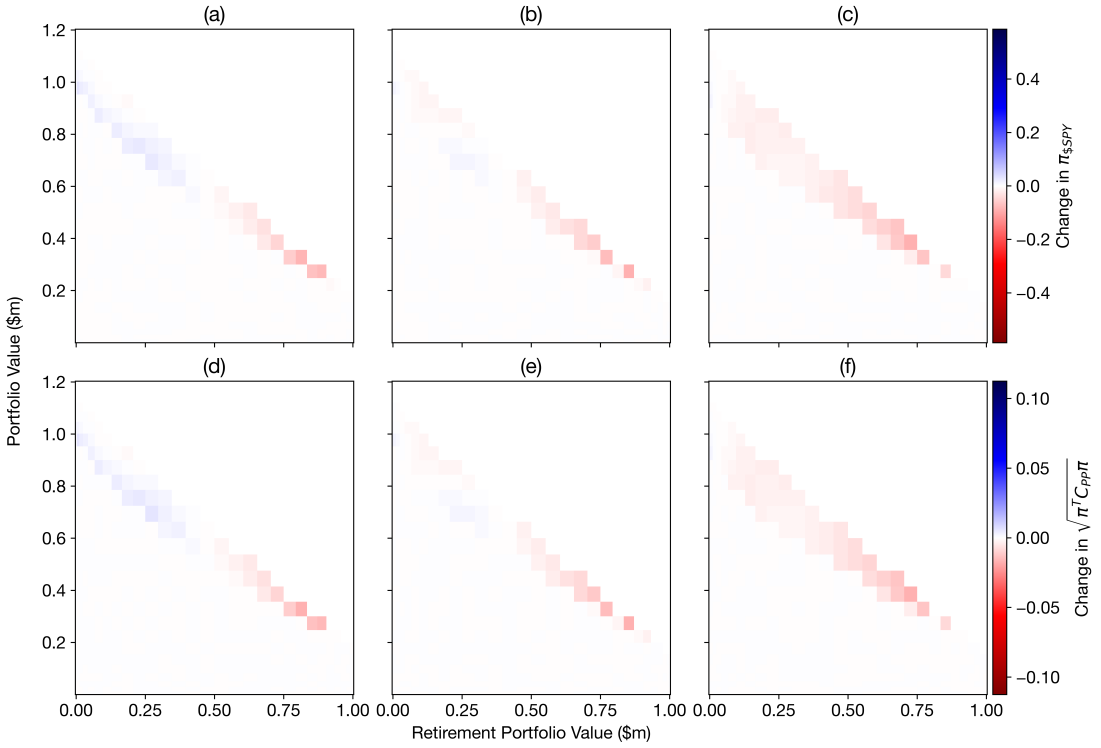

Figure 2.15: Change in Optimal Portfolio with Zero Employer Contribution, One Year Before Retirement per Income Level
Rows: Change in $\pi_{\$\text{SPY}}$; Change in $\sqrt{\pi^T C_{\text{PP}} \pi}$
Columns: $X_I \approx \$55{,}000$; $X_I \approx \$110{,}000$; $X_I \approx \$200{,}000$

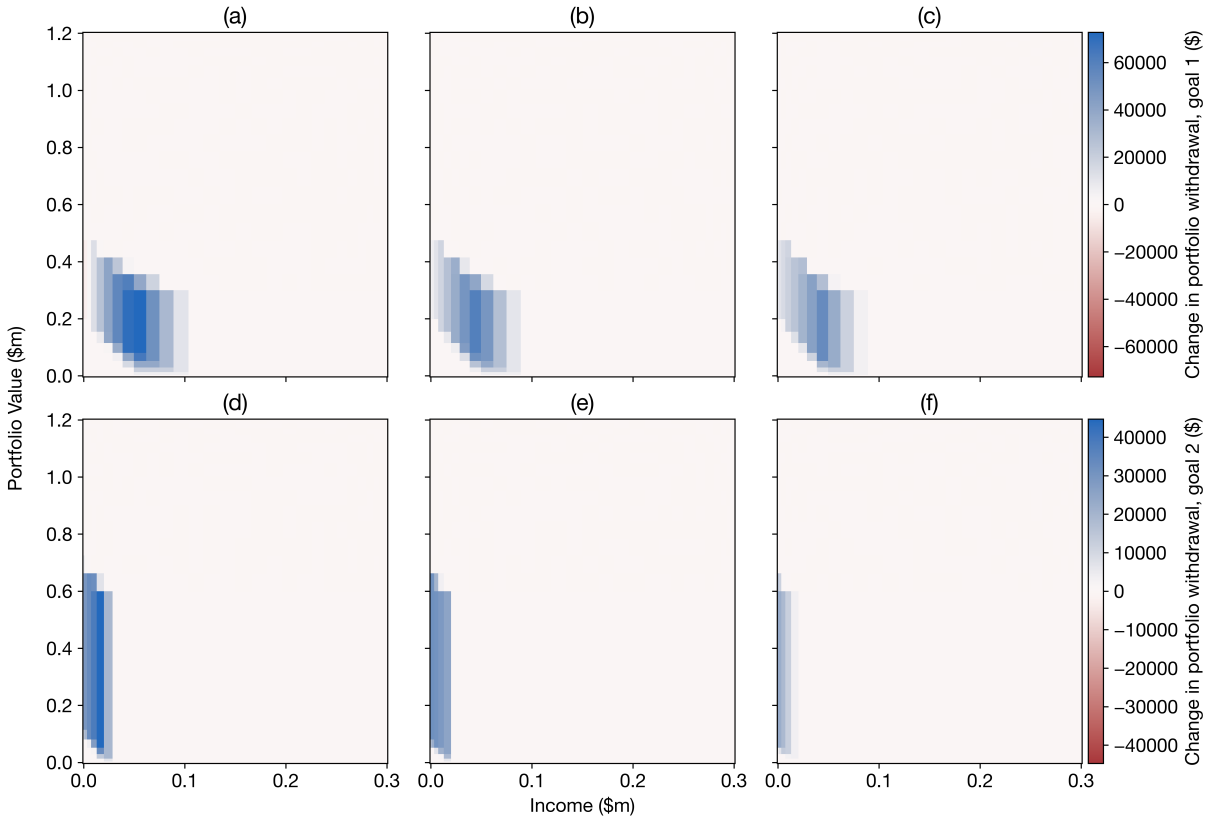

Figure 2.16: Difference Between Goal Contributions: Increased Income Growth Minus Baseline
Rows: Change in $\bar{G}_1$; Change in $\bar{G}_2$
Columns: $X_R \approx \$180{,}000$; $X_R \approx \$580{,}000$; $X_R = \$1{,}000{,}000$

2.4.4 Increased Income Growth Rate

It is expected that the increased income growth rate will have only positive implications for the investor's goals. This is demonstrated below.

2.4.4.1 Goal Contributions

We see that goal contributions are no less than those of the baseline investor, as per Figure 2.16. The investor still limits goal contributions for sufficiently low-X_P and low-X_I states, but to a lesser degree in general.

Figure 2.17 below shows the increase in $\zeta(0)$ for the investor with a higher income growth rate. An investor beginning their career with zero net worth may be expected to obtain

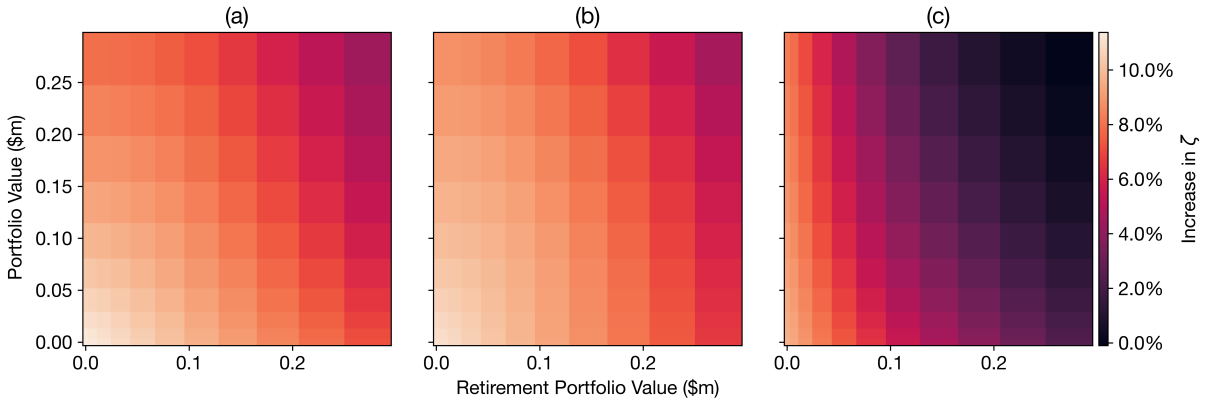

Figure 2.17: Difference in Initial Value Function: Increased Income Growth Minus Baseline (a): $X_I \approx \$55{,}000$; (b): $X_I \approx \$110{,}000$; (c) $X_I \approx \$200{,}000$

more than 10% extra utility than their counterpart with a lower income growth rate, for some initial income values. The advantage diminishes, however, as any of the initial state values increases.

2.4.4.2 Comparison of Safety Levels

Figure 2.18 below shows how changing α can affect the α-safe states. Figure 2.18(a), with a relatively high α of 0.05, eliminates important portfolio suggestions made at smaller α levels, as in Figure 2.18(b). However, too small an α can reveal the entirety of the state space, where the plurality of optimal solutions and polynomial approximation artifacts render irrelevant or unhelpful suggestions. A value of $\alpha = 0.01$ is therefore chosen for this analysis.

2.5 Conclusions and Future Work

In this research, we introduce a novel approach to GBWM for an investor with a retirement account and stochastic income. Our approach optimally determines the portfolio selection, income allocation, and goal contributions for an investor, taking into account their consumption preferences and financial situation. We conduct a series of numerical experiments to illustrate the flexibility and intuitiveness of our model. We show qualitative portfolio changes

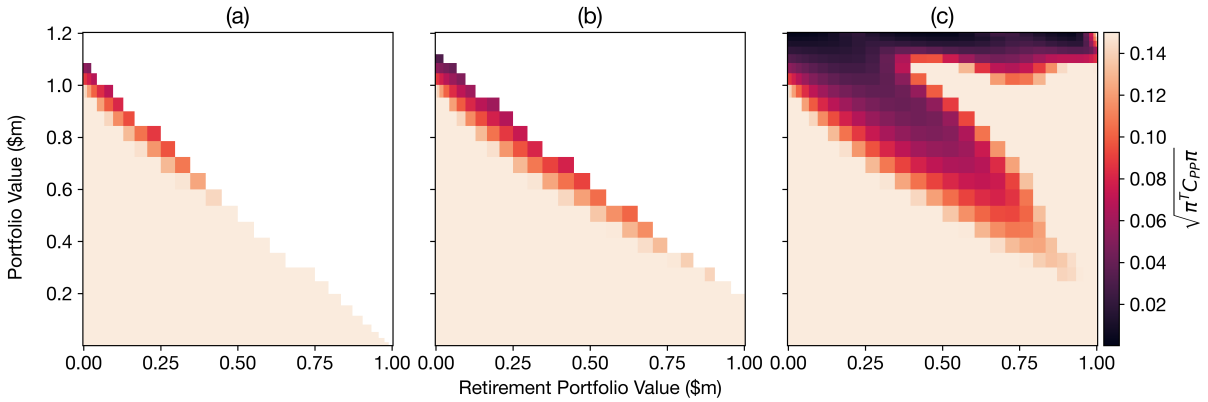

Figure 2.18: Baseline Investor Portfolio Volatility One Year Before Retirement per α
$X_I \approx \$55{,}000$
Columns: $\alpha = 0.05$; $\alpha = 0.01$; $\alpha = 0.001$

leading up to retirement with respect to a variety of factors. These include not only the investor's portfolio, retirement balance, and income, but also income-related considerations such as the employer's retirement contribution and the income's correlation to the market. Additionally, we analyze how the investor should allocate their income during their career.

Within the robo-advising industry, our model can be applied within both semi-autonomous and fully autonomous contexts. This stand-alone model facilitates automated portfolio allocations and suggests income allocations without the need for costly interactions with investment professionals. Within a financial advisor's toolkit, the model can provide interpretability of recommendations. Portfolio and income allocation suggestions can be compared via the derivatives of the value function to contextualize potentially unintuitive suggestions. Goal contributions, likewise, can be justified quantitatively with respect to expected future utility. Should investors wish to override the model's recommendations, a financial advisor can help them choose a suitable level of deviation from optimality to satisfy the investor's desires. If an investor's priorities have changed significantly at any point, the model may be solved again with an updated parameterization.

Our approach can be enhanced through various improvements. Many nuances of real-world investment are not accounted for under our simplified model. For tractability, our

continuous-time approach uses a simplified representation of the market, including proportional taxes, deterministic interest rates, simple asset and income dynamics, and a single retirement asset to invest in. Future research can always chase increased realism by relaxing or eliminating such assumptions. Beyond market-based assumptions, the client's goal deadlines and targets may not be deterministic. They may also be vulnerable to layoff concerns, particularly in poor market conditions, which our approach cannot capture fully. Employing methods from reinforcement learning may be a natural continuation to handle such nuances that, within a similar approach to this study, evade easy representation or computational tractability.

Computational improvements may be made to balance the problem resolution and runtime. Knowledge of the value function supremum and disengaged investor's value may help improve convergence if included as constraints in the Chebyshev polynomial fitting. Using complete Chebyshev polynomials for fewer parameters in the polynomial fitting may also provide sufficient precision with a reduction in solution time.

CHAPTER 3

Deep Reinforcement Learning for Goals-Based Wealth Management Under Random Goal Times

3.1 Introduction

Reinforcement learning methods have proven invaluable for solving challenging sequential decision-making problems. Difficulties that commonly hamper dynamic programming approaches, such as the curse of dimensionality and complicated environments, can be ameliorated with reinforcement learning. GBWM is a natural application area of reinforcement learning given the complexity of not only the market dynamics but also the investor's desires. Indeed, while the model in Chapter 2 presents useful themes of asset allocation, income distribution, and goal contributions, stochastic dynamic programming approaches may less comfortably deal with evolving market conditions than reinforcement learning methods. Applying reinforcement learning methods may therefore be a natural continuation of GBWM practice.

The effects of recessions are an important concern for investors. Involvement with financial planners during the GFC has been shown to help investors follow disciplined strategies for investment and portfolio rebalancing [59]. The value of robo-advising in helping clients manage the difficulties of economic downturns is thus clear, motivating the development of

GBWM algorithms that are recession-aware.

In addition to reductions in portfolio value and income, shocks can result in delayed purchases of durable goods [58]. A sensible question, therefore, is how an investors' decisions may change if their goal times are uncertain. Many purchases can be reasonably described as stochastic in time, such as the replacement of an aging vehicle, a relocation due to work or family, or a medical expense.

We consider a client with multiple competing goals, the purchase times of which are potentially uncertain. The client expects to make capital outlays around certain times, but is unsure exactly when. We include an economic variable to compare decision-making in recessions with normal times. Recommendations for goal contributions and portfolio allocations are made. As in Chapter 2, we conduct comparative experiments to evaluate how decision-making changes per the investor's goals and economic situation.

Our approach is to use reinforcement learning to determine optimal portfolio allocations between goals, with a static optimization at revealed goal times to recommend goal contributions and estimate value-to-go. Historical market data from recessionary and non-recessionary periods are used to sample risky asset returns. Our results show a sensible adaptation of investor behavior to their financial situation and environment. The potential of earlier goal times forces more conservative portfolio allocations. Goal contributions are greater for earlier goal times, and are intuitively balanced by the relative importance of the competing goals. The insights derived from this model, complementary to those in Chapter 2, support the development of a broader GBWM reinforcement learning framework in future.

3.2 Related Works

We provide a brief description of applications of reinforcement learning approaches to GBWM; broader context on GBWM is provided in Section 2.2. The viability of reinforcement learning methods is demonstrated by Das and Varma's work applying Q-learning to

obtain the same results as a previous GBWM dynamic programming approach [29]. Dixon and Halperin use G-learning, a probabilistic extension of Q-learning, to optimize a portfolio for an investor who contributes to their portfolio during employment before periodically withdrawing post-retirement [34]. Bauman et al. use a deep reinforcement learning approach for a single all-or-nothing investment goal [10].

Multiple previous works examine portfolio allocation under an uncertain terminal time, typically within the context of retirement or death. Yaari examines optimal consumption under an uncertain death time [100]. Further work introduces risky assets and portfolio allocation to similar problems [46, 47, 69]. Blanchet-Scalliet et al. solve the optimal investment problem in which the time horizon is dependent on the returns of risky assets [13].

To our knowledge, this work is the first application of reinforcement learning to multiple competing goals, stochastic goal times, and recessionary dynamics. The rest of this paper is structured as follows. Section 3.3 describes the market and investor's goals, formulating the value function and optimal goal contribution equations. Section 3.4 describes the reinforcement learning approach used to solve these equations and the numerical experiments conducted. Results are presented in Section 3.5, followed by conclusions in Section 3.6.

3.3 Problem Description

3.3.1 The Market and Investor

The investor has a portfolio value at time t of $X(t)$, allocated between a risky asset, with price $S(t)$, and a risk-free asset, with price $B(t)$. The risk-free asset earns interest at a rate $r \geq 0$, yielding the dynamics

$$dB(t) = rB(t)\,dt. \tag{3.1}$$

The market is subject to infrequent recessions. The economic state is represented by a binary variable,

$$E(t) = \begin{cases} 0 & \text{market is not in a recession} \\ 1 & \text{market is in a recession} \end{cases} \tag{3.2}$$

which evolves according to a continuous-time Markov chain with (regular) rate matrix, $Q : \{0,1\}^2 \to \mathbb{R}$. The corresponding transition probability matrices for durations $t > 0$ are denoted by $P(t) = e^{tQ}$.

The stock is assumed to evolve according to a Itô process characterized by the economic state,

$$dS(t) = \mu(t, E)S(t)\,dt + \sigma(t, E)S(t)dW(t) \tag{3.3}$$

where W is a Brownian motion, μ is the stock's drift, and σ is the volatility of the stock. The stock drift and volatility are jointly distributed according to the economic state:

$$\mu(t, E), \sigma(t, E) \sim f_{(\mu,\sigma)|E}(t) \tag{3.4}$$

We denote by $\mathcal{F}_t^W$ and $\mathcal{F}_t^E$ the filtrations generated by W and the Markov chain, respectively, by time t. We assume these filtrations are independent, and denote by $\mathcal{F}_t \triangleq \mathcal{F}_t^W \vee \mathcal{F}_t^E$ their join.

The investor's income is dependent on the economic regime. Normally, they earn I_0 dollars per year, but in a recession this decreases to αI_0 for some $\alpha \in [0, 1]$. In line with Chapter 2, we assume the investor's income tax rate is 15% and that they spend 50% of net income on necessities, thus ultimately contributing 42.5% of gross income to the portfolio.

The portfolio dynamics are therefore

$$\begin{aligned} dX(t,E(t),\pi(t)) = &\left[(\pi(t)\mu(t,E(t)) + (1-\pi(t))r)\,X(t) + I(t,E(t))\right]\,dt \\ &+ \pi(t)\sigma(t,E(t))X(t)\,dW(t) \end{aligned} \tag{3.5}$$

$$I(t,E(t)) = \begin{cases} 0.425I_0 & E(t)=0 \\ 0.425\alpha I_0 & E(t)=1 \end{cases} \tag{3.6}$$

3.3.1.1 Goals

The formulation of goals is similar to that in Section 2.3.1.2, with the additional element of goal-time uncertainty. The investor has K consumption goals, occurring at times τ_k, $k = 1, 2, ..., K$. At each goal, the investor withdraws $G_k(\tau_k, X(\tau_k), E(\tau_k))$ from their portfolio, earning some utility $u_k(G_k(\tau_k, X(\tau_k), E(\tau_k)))$ from this goal. The utility functions are assumed to be bounded and concave, with $\frac{\partial u_k}{\partial G_k} \to 0$ as $G_k \to \infty$.

Goal times occur randomly, with each goal having a deterministic probability distribution over the time domain of $f_k(t)$, $t \in [0, T]$. That is, $(\tau_k)_{k\in\{1,2,...,K\}}$ is a sequence of stopping times independent of the filtration $\mathcal{F}_t$. In the case of $f_k(t) = \delta(x - t_k)$ for $t_k \in [0, T]$, this simplifies to a deterministic goal time. For convention, let $\tau_0 = 0$. We assume prior knowledge of the order of the coming goals and that no two goals will occur concurrently. We therefore have $(f_1, f_2, ..., f_K)$ such that $\mathbb{P}(\tau_1 < \tau_2 < ... < \tau_K) = 1$. We may therefore isolate the boundaries of the goal times as $\tau_k^- = \inf \text{supp} f_k$ and $\tau_k^+ = \sup \text{supp} f_k$, thereby obtaining

$$\tau_k^- \leq \tau_k^+ < \tau_{k+1}^- \leq \tau_{k+1}^+, \quad k = 1, ..., K-1 \tag{3.7}$$

with $t_0^- = 0$ for convention. We make the natural assumption that the distribution of each goal time has bounded support, i.e. $\sup \text{supp} f_k < \infty \; \forall k$.

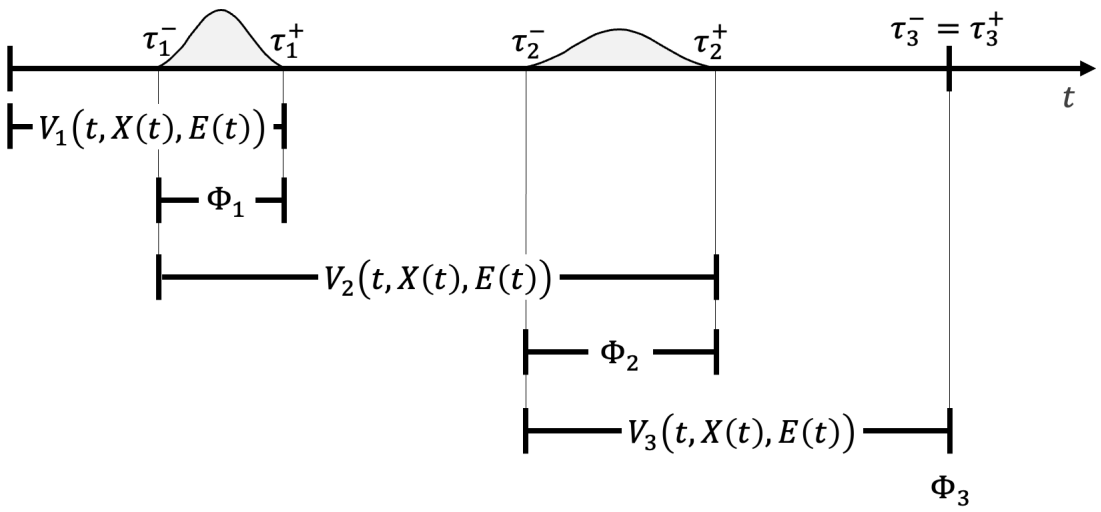

Figure 3.1: Schematic Division of a 3-Goal Time Domain into Separate Value Functions Under Uncertain Goal Times

3.3.2 Value Function

Using the above notation, we define the series of value functions for $k = 1, 2, ..., K$ as follows:

$$V_k(t, X(t), E(t)) = \mathbb{E}^{\pi^\star}\left[\int_t^{\tau_k^+} f_{k|t}(s)\Phi_k(s, X(s), E(s))\, ds \;\middle|\; \mathcal{F}_t\right], \quad \tau_{k-1}^- \le t \le \tau_k^+ \tag{3.8}$$

$$\begin{aligned}\Phi_k(t, X(t), E(t)) = &\max_{G_k \in [0, X(t)]} \{u_k(G_k(t, X(t), E(t))) + V_{k+1}(t, X(t) - G_k(t, X(t)), E(t))\} \\ &\tau_{k-1}^- \le t \le \tau_k^+\end{aligned} \tag{3.9}$$

where $\mathbb{E}^{\pi^\star}$ denotes expectation under the optimal action, and $f_{k|t}(s)$ is the conditional distribution of $\tau_k = s$ given $\tau_k > t$. For convention, we have that $V_{K+1} = 0$. Figure 3.1 demonstrates how the value functions and terminal conditions are defined over an example problem with two uncertain goals and one deterministic goal.

The concavity of the value function in X can easily be seen via a similar inductive argument to that in Theorem 1, as outlined in Theorem 5 below.

Theorem 5. *$V_k(t, x, E)$ and $\Phi_k(t, x, E)$, for $k = 1, 2, ..., K$, $E \in \{0, 1\}$, and $t \in \left[\tau_{k-1}^-, \tau_k^+\right]$*

are concave in x.

Proof. Trivially, $V_{K+1} = 0$ is concave in x. Suppose V_{k+1} is concave, for some $k = 1, 2, ..., K$ and all $t \in \left[\tau_k^-, \tau_k^+\right]$ and $E \in \{0, 1\}$. For $j = 1, 2$, let $x^{(j)}$ be arbitrary feasible portfolio values at time $t \in \left[\tau_k^-, \tau_k^+\right]$ with respective optimal goal selections $G_k^1(t)$ and $G_k^2(t)$ such that, for $\eta \in [0, 1]$, $x^{(3)} \triangleq \eta x^{(1)} + (1-\eta)x^{(2)}$. We have that $G_k^{(3)}(t) \triangleq \eta G_k^{(1)}(t) + (1-\eta)G_k^{(2)}(t)$ is a feasible contribution for $x^{(3)}$ as

$$\eta G_k^{(1)}(t) + (1-\eta)G_k^{(2)}(t) \leq (1-\nu_P)(\eta x^{(1)} + (1-\eta)x^{(2)}) \tag{3.10}$$

$$= (1-\nu_P)x^{(3)}. \tag{3.11}$$

Therefore,

$$\begin{aligned}
&\eta\Phi_k(t, x^{(1)}, E) + (1-\eta)\Phi_k(t, x^{(2)}, E) \\
&\quad = \eta u_k(G_k^{(1)}(t)) + (1-\eta)u_k(G_k^{(2)}(t)) \qquad (3.12)\\
&\qquad + \eta V_{k+1}\left(t, x^{(1)} - G_k^{(1)}(t), E\right) + (1-\eta)V_{k+1}\left(t, x^{(2)} - G_k^{(2)}(t), E\right) \\
&\quad \leq u_k(G_k^{(3)}(t)) + V_{k+1}\left(t, x^{(3)} - G_k^{(3)}(t), E\right) \qquad (3.13)\\
&\quad \leq \Phi_k(t, x^{(3)}, E), \qquad (3.14)
\end{aligned}$$

for $E \in \{0, 1\}$, and thus Φ_k is concave in x.

Now, for fixed $t \in [\tau_{k-1}^-, \tau_k^+]$ and any $E(t) \in \{0, 1\}$, let $\pi^{(i)}$, $i \in \{1, 2\}$, be optimal portfolios for $X^{(1)}(t) = x^{(1)}$ and $X^{(2)}(t) = x^{(2)}$, respectively, over $[t, \tau_k^+]$. Trivially,

$$\pi^{(3)} \triangleq \frac{\eta\pi^{(1)}X^{(1)} + (1-\eta)\pi^{(2)}X^{(2)}}{\eta X^{(1)} + (1-\eta)X^{(2)}} \tag{3.15}$$

is an admissible portfolio and $x^{(3)} \triangleq \eta x^{(1)} + (1-\eta)x^{(2)}$ is a feasible state. Following a similar argument to Theorem 1, it is easily seen that $X^{(3)}(s) = \eta X^{(1)}(s) + (1-\eta)X^{(2)}(s)$ for $t \leq s \leq \tau_k$ for any $\tau_k \in \left[\max\left\{t, \tau_k^-\right\}, \tau_k^+\right]$. Because taking the expectation of Φ_k over

potential goal times preserves concavity, the concavity of the value function is readily seen [14]. □

3.4 Methods

This section details the reinforcement learning approach used to solve the problem. Relevant reinforcement learning background is provided in Section 3.4.1 before the problem and solution descriptions are given in Sections 3.4.2 through 3.4.4.

3.4.1 Reinforcement Learning Preliminaries

To contextualize the following approach, we must first cover relevant topics in reinforcement learning and how they relate to the problem at hand. Reinforcement learning is the theory and practice of learning an agent's optimal policies for interacting with their environment to maximize their reward. Unlike dynamic programming, reinforcement learning does not compute the optimal policy and value function, instead learning the correct decisions from trial and error.

3.4.1.1 Markov Decision Process

The Markov decision process (MDP) formalizes the agent-environment interaction and is therefore key to reinforcement learning [84]. An MDP consists of the tuple (S, A, P_a, R_a), where: S is the state space; A is the action space; $P_a(s, s')$ is the probability of transitioning into state s' from s by taking action a; and $R_a(s, s')$ is the immediate reward received after the aforementioned transition. We model this GBWM problem as a series of reinforcement learning subproblems, concatenated via solving continuous optimization problems for each goal. Each subproblem is constructed as a MDP with which an episodic reinforcement learning agent is trained.

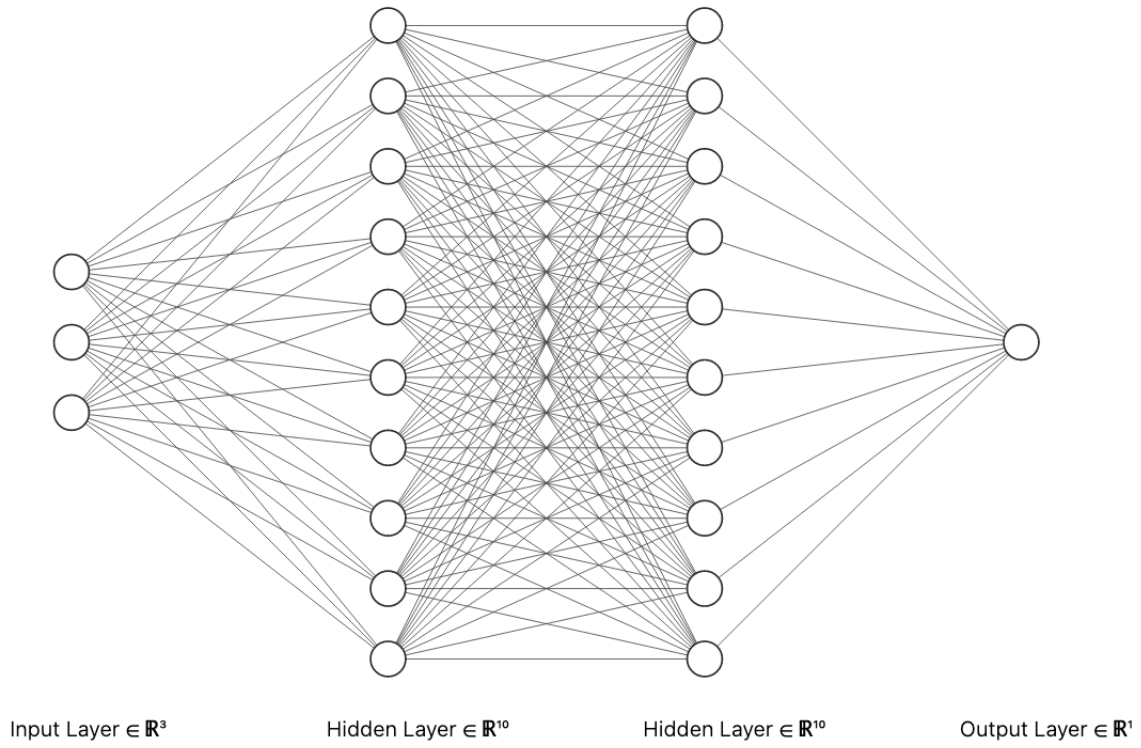

Figure 3.2: Artificial Neural Network Schematic [60]

3.4.1.2 Function Approximation and Neural Networks

In our GBWM context, the state is continuous with respect to both time and the portfolio value but discrete with respect to the economic state. The action space, i.e. portfolio allocation, is continuous. While tabular solution methods are effective for small state and action spaces, the infinitely large spaces created by continuous states and actions render these methods infeasible [84]. Instead, function approximation via *artificial neural networks* is used to provide continuous mappings between states and values or policies. Artificial neural networks are functions composed in a structure inspired by the human brain. An input layer receives data which is subsequently propagated and transformed through hidden layer(s) and delivered to the output layer. These transformations are determined by activation functions parameterized by values which are trained to approximate a function of choice. Figure 3.2 shows an example artificial neural network architecture. An artificial neural network with multiple hidden layers is said to be a *deep neural network*.

Deep reinforcement learning is the application of deep neural networks to approximate the value function, policy function, or both. Deep reinforcement learning methods that approximate both the value and policy functions are known as actor-critic methods. We apply one such algorithm, Proximal Policy Optimization (PPO), which has been previously applied in an GBWM context [77, 10].

3.4.2 Problem and Solution Setup

We solve a series of problems to compare optimal portfolio allocation across multiple investment scenarios. Each problem consists of K reinforcement learning subproblems that are linked via (3.9). At the beginning of each episode, the goal time is selected according to the goal-time distribution. The episode terminates when the time step reaches the goal time. At prior time steps, the investor decides their portfolio allocation not yet knowing whether the economy will be in recession for the coming year. The economy and portfolio are then sampled for the next year.

Similarly to Chapter 2, we explore the problem parameterization around a "baseline" investor to assess the impact of economic and goal scenarios on their decision-making. The following subsections detail these experiments.

3.4.2.1 The Baseline Investor

The baseline investor has two goals with targets of \$250,000 and \$500,000, respectively. The first goal is stochastic, being equally likely to occur at the beginning of years 8 through 12. The second goal is deterministic and will occur at year 20. As probability mass functions,

these are expressed as

$$\mathbb{P}(\tau_1 = t) = \begin{cases} \frac{1}{5} & t \in \{8, 9, 10, 11, 12\} \\ 0 & \text{otherwise} \end{cases} \tag{3.16}$$

$$\mathbb{P}(\tau_2 = 20) = 1 \tag{3.17}$$

Each goal is considered equally important by the investor. Taking a similar utility function to that used in Section 2.3.4.2, our utility functions here are specified as

$$u_1(x) = \min\left\{\frac{x}{\varphi_1}, 1\right\} \tag{3.18}$$

$$u_2(x) = \min\left\{\frac{x}{\varphi_2}, 1\right\} \tag{3.19}$$

with $(\varphi_1, \varphi_2) = (250000, 500000)$. These are shown in Figure 3.3 below.

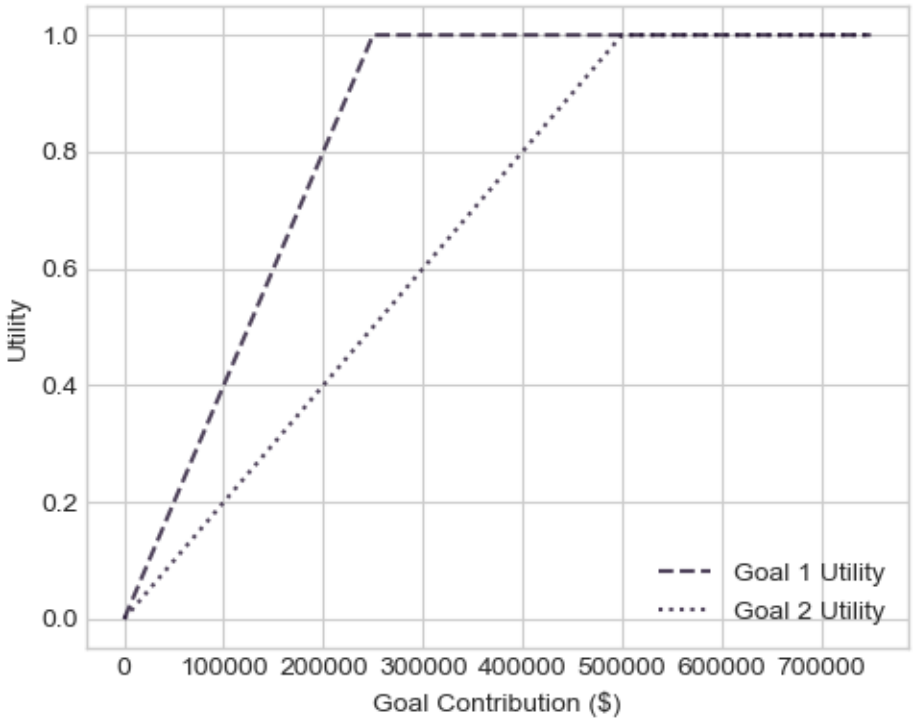

Figure 3.3: Utility Functions of the Baseline Investor

The investor's gross income in non-recessionary times is assumed to be \$60,000 annually. We take $\alpha = 1 - 0.042$, in line with the 4.2% decrease in median decrease in US household incomes in the 2007-2009 period [55]. The risk-free rate of return is set to 3%.

3.4.2.2 Deterministic Goal Times

We consider an investor certain of their goal times to assess the impact of goal-time uncertainty on portfolio decisions. We set the goal time of the first goal to be the latest possible time, i.e.

$$\mathbb{P}(\tau_1 = 12) = 1. \tag{3.20}$$

3.4.2.3 Different Goal Priorities

The relative weightings of the goals importantly influences goal contributions. We consider the case where the second goal is more important than the first, decreasing the range of u_1:

$$u_1(x) = \frac{1}{3} \min \left\{ \frac{x}{\varphi_1}, 1 \right\} \tag{3.21}$$

3.4.3 The Markov Decision Process

3.4.3.1 State Space

The state for each reinforcement learning problem consists of components representing time, the portfolio value, and the economic state. We choose to represent the portfolio value state as the ratio of the portfolio value to the "safe level" of portfolio value for which all future utility can be satisfied with probability 1. Unlike in Chapter 2, this formulation admits a finite safe portfolio level. This approach clarifies how decision-making changes approaching safe wealth levels.

The safe level in this problem refers to the minimum portfolio value at any time that, by investing exclusively in the risk-free asset thereon, will fully satisfy all future goals even

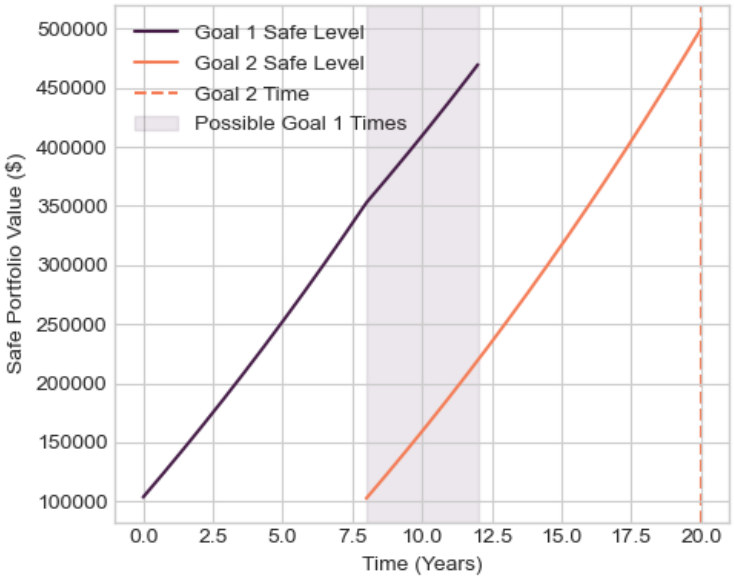

Figure 3.4: Safe Levels of Portfolio Value for the Baseline Investor

under the worst-case income scenario. We perform a change of variable similar to that of Capponi and Zhang (2022) [21]. The safe level for goal $k = 1, 2, ..., K$ is defined as

$$\tilde{X}_k(t) = \begin{cases} e^{-r(\tau_k^- - t)}(g_k + \tilde{X}_{k+1}(\tau_k^-)) - \frac{(1-\nu_I)\alpha I_0}{r}(1 - e^{-r(\tau_k^- - t)}) & t < \tau_k^- \\ g_k + \tilde{X}_{k+1}(t) & t \in [\tau_k^-, \tau_k^+] \end{cases} \tag{3.22}$$

for $r > 0$, $t \in [\tau_{k-1}^-, \tau_{k+1}^-]$, and $\tilde{X}_{K+1} = 0$. In the case where $r = 0$, we have

$$\tilde{X}_k(t) = \begin{cases} g_k + \tilde{X}_{k+1}(\tau_k^-) - (1 - \nu_I)\alpha I_0(\tau_k^- - t) & t < \tau_k^- \\ g_k + \tilde{X}_{k+1}(t) & t \in [\tau_k^-, \tau_k^+] \end{cases} \tag{3.23}$$

Figure 3.4 shows the safe levels for the baseline investor.

The state for problem k at time $t \in [\tau_{k-1}^-, \tau_k^+)$ is therefore defined as $s_t^k = \left(\frac{t-\tau_{k-1}^-}{\tau_k^+ - \tau_{k-1}^-}, \frac{X}{\tilde{X}_k(t)}, E(t)\right) \in [0, 1) \times [0, \infty) \times \{0, 1\}$. Although the gradient of the safe level around τ_1^- is roughly the same, an uncertain terminal goal time would have a noticeable gradient discontinuity at τ_K^- as the safe value becomes constant at g_K. We leave to future

work an assessment of the comparative strengths of using the safe level as the portfolio state versus, say, a scaled portfolio value.

3.4.3.2 Action Space

The action space is equal across all goals and times, consisting solely of the portfolio, $\pi(t) \in [0, 1]$.

3.4.3.3 State Transitions

The economic state transition matrix and the risky asset dynamics parameters are informed by historical returns of the SPDR S&P 500 ETF ($SPY) per U.S. recession dates inferred by a GDP-based recession indicator between 2000 and 2022 [48]. We estimate the year-to-year transition probability matrix using the frequency of quarterly transitions between recessionary and non-recessionary economies.

For the risky asset dynamics, we use a similar sampling approach to Bauman et al. [10]. For each episode time step, in a (non-)recessionary economy, we sample twelve of the monthly $SPY log-returns with replacement from the (non-)recessionary quarters within the dataset. These are used to construct the drift and volatility terms in (3.3), from which a sample yearly return is taken.

3.4.3.4 Reward Function

Rewards are nonzero only at the goal time. For the last goal, the entire portfolio is liquidated, so $R_a(s, (1, X, E)) = u_k(X)$ for all a and s. For in-between goals, the terminal reward is determined by balancing the current goal's utility with future goals via an estimate of (3.9). The approach used is described in Section 3.4.4.2.

3.4.4 Procedure

The problem is solved via recursion. One reinforcement learning problem is solved per goal to estimate the value function and optimal policies, detailed in Section 3.4.4.1. The optimal goal contributions in-between are estimated with a convex optimization approach, specified in Section 3.4.4.2.

3.4.4.1 Algorithm

The PPO actor and critic neural networks each have 2 hidden layers with 10 neurons each. This architecture is shown in Figure 3.2 above. The hyperbolic tangent activation function is used. The batch size used is 32,768. The learning rate used is 0.001. Each subproblem is trained for 80,000,000 total timesteps. Implementation was conducted with Stable Baselines3 and Pytorch [74, 72].

The initial state is sampled as follows. The initial time is sampled uniformly from the prior goal's possible goal times, to allow for exploration of states with low portfolio values at later times. The initial safety value is sampled uniformly within $[0, 1]$. The economic state is sampled from its stationary distribution.

3.4.4.2 Optimizing the Goal Contributions

Rewards are only nonzero at goal times. Except for the final goal, the reward is made by balancing the current goal's utility and expected future utility. As the argument of the max function is concave, the utility-maximizing goal contribution is tractable in theory via convex optimization as in Section 2.3.2.1. However, the value function estimate from the reinforcement learning algorithm may not maintain concavity across the entire domain and, in practice, are occasionally only quasiconcave. Another challenge posed is that, even if the value function is concave, each rollout may involve solving a computationally expensive convex optimization problem for the optimal goal contribution. We address both of these problems by fitting a globally concave finite difference approximation of the value function

with which we construct estimates of the optimal contributions.

Trivially, the optimal contribution for the last goal is the investor's entire portfolio. For problem $k < K$, let $t \in [\tau_k^-, \tau_k^+]$. Denote by $V_{k+1}^{\boldsymbol{\theta}}(t, X, E)$ the value function estimate at this interval for portfolio value X and economic state E, where $\boldsymbol{\theta}$ represents the parameters of the critic neural network. We fit a concave piecewise linear approximation of the value function after the goal contribution for each economic state, $j \in \{0, 1\}$. A grid of $0 = X_0 < X_1 < ... < X_N$ is used, where X_N is the safe portfolio level after the goal contribution. The approximation is calculated with

$$\begin{aligned} \underset{y_i, i=1,2,...,N}{\text{minimize}} \quad & \sum_{i=0}^{N} \left\| \sum_{\tilde{j}=0}^{1} P_{j,\tilde{j}} V_{k+1}^{\boldsymbol{\theta}}(t, X_i, \tilde{j}) - y_{i,j} \right\|_2^2 \\ \text{s.t.} \quad & y_{i,j} - y_{i-1,j} \geq 0, \quad i = 2, ..., N \\ & y_{i,j} - 2y_{i-1,j} + y_{i-2,j} \leq 0, \quad i = 3, ..., N \end{aligned} \tag{3.24}$$

where the first set of conditions ensures the value function is nondecreasing and the second set ensures concavity. Here, $P_{j,\tilde{j}}$ represents the transition probability matrix between states j and $\tilde{j}$. Denote by $\hat{V}_{k+1}(t, x, j)$ the concave piecewise linear interpolation of the $y_{i,j}$ values. The continuity of the terminal value function is guaranteed by its concavity. For the grid points $0 = b_0 < b_1 < ... < b_N = X_N + \varphi_k$, we take

$$\phi_i(t, j) = \max_{g \in [0, b_i]} \left\{ u_k(g) + \hat{V}_{k+1}(t, b_i - g, \tilde{j}) \right\}, \quad i = 1, ..., N, \quad j = 0, 1 \tag{3.25}$$

to be the corresponding terminal values. A piecewise linear interpolation is then applied for each economic state which we denote by $\hat{\Phi}_k(t, X, j)$. This interpolation serves as an efficient way of estimating terminal value utility. We note that due to the piecewise linear form of (3.25), the maximization can be performed via solving a linear program.

Denote by $\mathbf{G}_{k,j}(x)$ the feasible set of contributions at goal k for portfolio value $x \in \mathbb{R}_{\geq 0}$ and

economic state j. Berge's maximum theorem implies that the maximizers of (3.25), denoted by the correspondences $\mathbf{G}^{\star}_{k,j}(x)$, $k = 1, ..., K$, $j = 0, 1$, are each upper hemicontinuous and have nonempty, compact values [71]. While (3.25) is not strictly concave, we find in practice that it produces unique maximizers aside from when $\exists g \in \mathbf{G}^{\star}_{k,j}(x)$ with $g > \varphi_k$. We therefore assume the functions $\hat{G}^{\star}_{k,j}(x) \triangleq \min \mathbf{G}^{\star}_{k,j}(x)$ are continuous, and take their piecewise linear interpolations at b_i, $i = 1, ..., N$, as our optimal goal contribution functions.

3.5 Results

We first present the baseline investor results, followed by the comparative studies. Common across all experiments is near-total investment in the risk-free asset during recessionary times. This is expected, given the negative expected return of the risky asset per the data used in Section 3.4.3.3. We therefore choose to show only the portfolios for non-recessionary states.

3.5.1 The Baseline Investor

Figures 3.5 and 3.6 below show the allocation to the risky asset leading up to the first and second goals, respectively, in terms of both portfolio value and portfolio safety. In both subproblems, the investor transitions from full investment in the risky asset to none as their portfolio value approaches the safe level.

Figure 3.7 below shows the contributions towards the first goal. For all goal times and economic states, the investor funds the goal as much as possible. This is a sensible policy given the high marginal utility of the first goal.

3.5.2 Different Goal Priorities

Figure 3.8 below shows the allocation prior to the first goal for the investor who weights the second goal higher than the first. Again, the allocation to the risky asset approaches zero as the portfolio value approaches safety. However, this investor maintains full investment

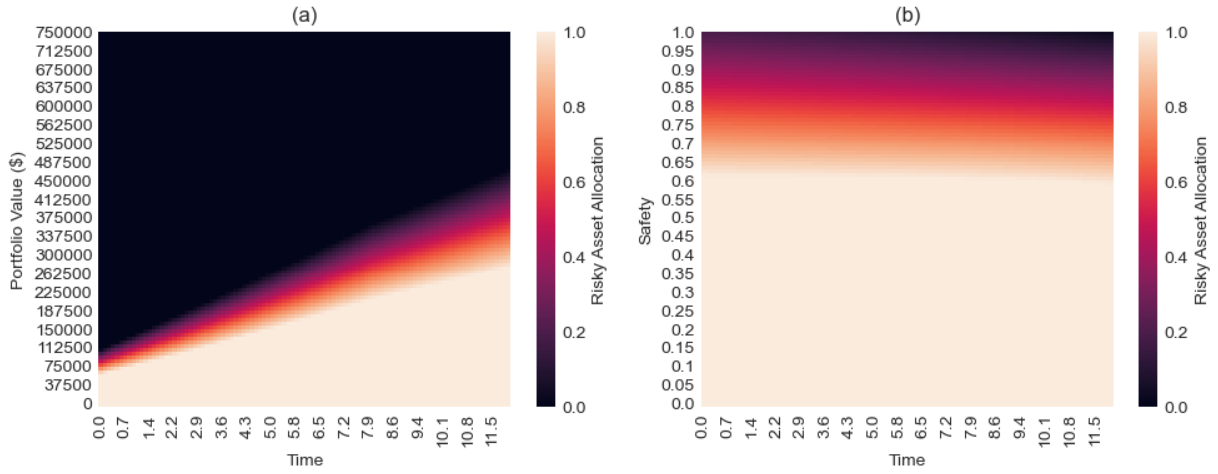

Figure 3.5: Baseline Investor: Portfolio Allocations Before Goal 1 by Portfolio (a) Value and (b) Safety

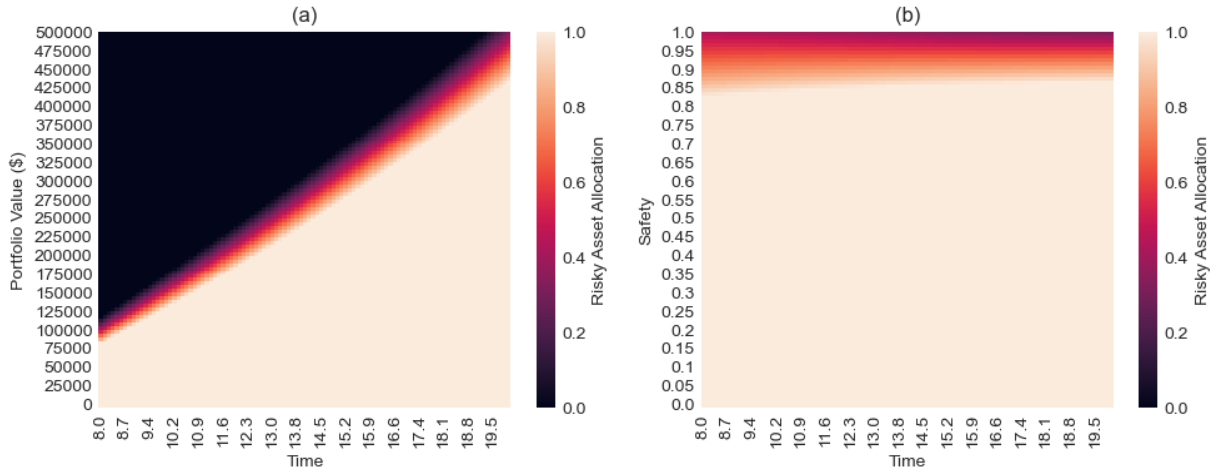

Figure 3.6: Baseline Investor: Portfolio Allocations Before Goal 2 by Portfolio (a) Value and (b) Safety

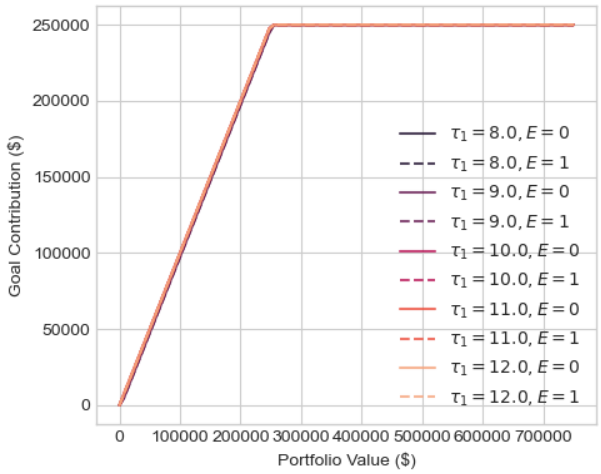

Figure 3.7: Baseline Investor: Contributions to Goal 1

in the risky asset up to a higher portfolio safety, as exemplified by image (b). Figure 3.9 below further demonstrates the relative importance of the second goal. Unlike Figure 3.7, we see that the investor never contributes the entirety of their portfolio balance to the first goal, opting instead to conserve funds towards the second goal. We also see that the investor contributes less to the goal for a later goal time. Intuitively, a later goal time leaves less time to accrue funds for the second goal. There is a minor influence of the economic state on the goal contributions: contributions are slightly decreased if the goal time occurs during a recession.

3.5.3 Certain Goal Times

Figure 3.10 below shows the optimal risky asset allocation leading up to the first goal when the goal occurs with certainty after twelve years. Near the goal time, the allocation is almost the same as that in Figure 3.5. However, when the goal time is known, the investor takes additional risk at higher safety levels and earlier times. Clearly, the possibility of an early goal time impels the investor to reduce risk earlier for some portfolio safety levels, lest an unfavorable market return and early goal time force a low contribution towards the goal.

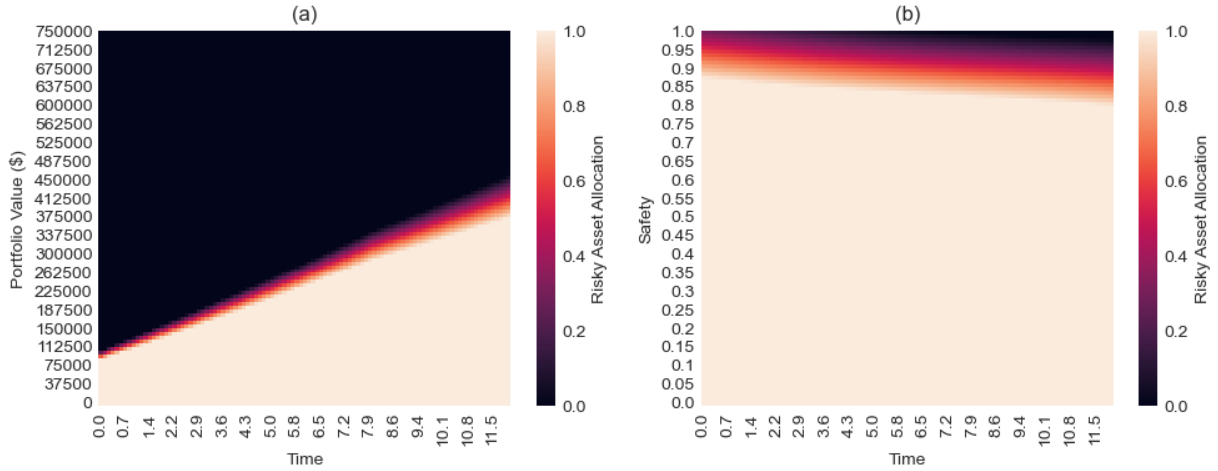

Figure 3.8: Different Goal Priorities Investor: Portfolio Allocations Before Goal 1 by Portfolio (a) Value and (b) Safety

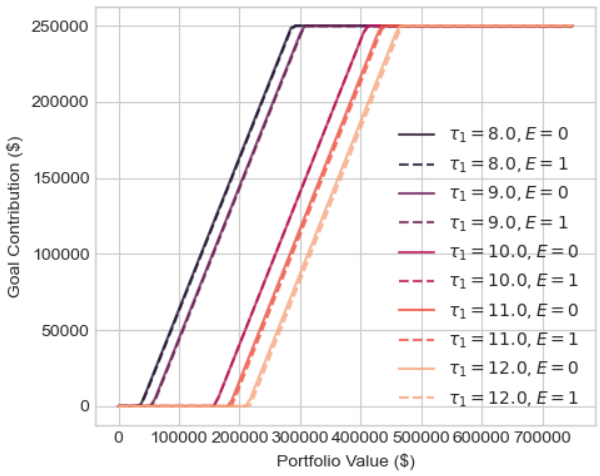

Figure 3.9: Different Goal Priorities Investor: Contributions to Goal 1

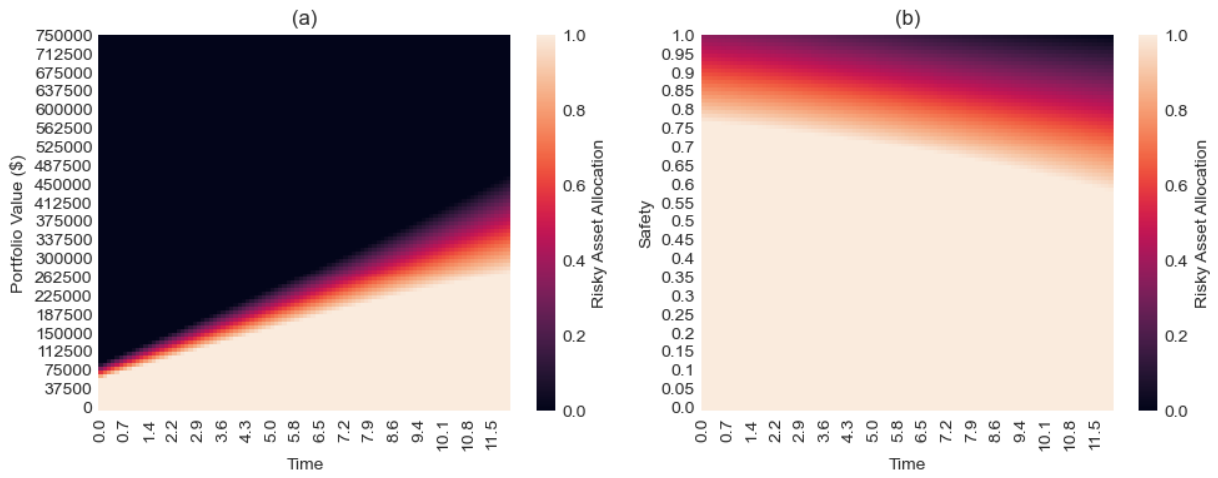

Figure 3.10: Certain Goal Times Investor: Portfolio Allocations Before Goal 1 by Portfolio (a) Value and (b) Safety

3.6 Conclusions

This work extends the reinforcement learning frontier of GBWM by optimizing decision-making of an investor with multiple competing goals with uncertain times. Our procedure estimates the optimal portfolio allocation and goal contributions based on an investor's consumption preferences and goal uncertainties. Our experiments demonstrate sensible portfolio recommendations based on the investor and their economic environment. Suggested goal contributions are intuitive.

There are a plurality of possible extensions to our approach. Many of the facets of the model in Chapter 2 can be adapted to this model: a retirement account with income contribution decisions, tax implications, and sophisticated income dynamics would be welcome additions. While treatment of the market and recession dynamics here is relatively simple, reinforcement learning is flexible to a wide description of market behavior. Future research may include a more nuanced economic simulator, including interest rate dynamics, unemployment considerations, and latent economic states. Furthermore, interaction between the goal time distribution and market behavior would be a worthwhile improvement. Lastly, a sensible extension would be to convert the sequential reinforcement learning problems into

a single problem, representing the goal contributions within the MDP state and action. Not only would this avoid needing direct optimization of the goal contributions (and therefore knowledge of the economic state transition probabilities), but also allow for nonconcave utility functions, such as those for all-or-nothing goals.

CHAPTER 4

External Forces on Financial Markets: Evidence from the GameStop Short Squeeze and Flash Crash

4.1 Introduction

Markets are highly complex and interdependent, influenced by a multitude of factors. Exogenous events and unique market phenomena can dramatically destabilize prices. Such abnormalities are difficult to predict and often have unprecedented consequences. In addition, it can be challenging to decompose market price action into endogenous components (e.g. changes in fundamental stock price valuations) and exogenous components (e.g. the market crash due to COVID-19) [8].

The GameStop short squeeze of January 2021 is a clear example of an exogenous event on market prices [88]. Retail investors on social news website Reddit instigated a rally of GameStop's stock price [61]. The stock saw a 1,500% increase in price over a two-week period ending January 27 [61]. A subsequent -44% crash of the stock price occurred the next day. A variety of other so-called *meme stocks* simultaneously experienced similar price behavior, including those of AMC Entertainment Holdings, Inc., Bed Bath & Beyond Inc., and Eastman Kodak Company [82].

Another example is the flash crash of May 6, 2010, which resulted in steep drawdowns and

recoveries of major stock indices within 36 minutes [54]. During, the Dow Jones Industrial Average sustained its second largest intraday point decline, dropping approximately 9% [54].

The idiosyncrasies of each abnormality make risk management and portfolio allocation difficult. The flash crash was a market-wide phenomenon, whereas the GameStop short squeeze was isolated to a few stocks. While there are many models in the literature connecting market abnormalities, such as the COVID-19 pandemic, to markets, most do so on the mesoscopic level at highest. That is, the behavior of a relevant, yet often small, subset of traded assets are examined under econophysical or econometric methods. In contrast, macroscopic models, which consider all traded assets in a market simultaneously, capture interdependencies between assets invisible to smaller-scale models. There are conceptual benefits for risk management in taking such a perspective.

This research addresses the aforementioned challenges by aiming to quantify external influences on markets macroscopically. To do so, we extend the AlShelahi and Saigal macroscopic model of equity markets, a physics-inspired model that treats each stock as a particle within an Eulerian fluid-flow system of stochastic partial differential equations [5]. Calibration of this model during the flash crash has indicated the model describes market abnormalities. Our contribution is the decomposition of stock price acceleration into an endogenous and exogenous component, the latter of which we term *investor impatience*. We compare these external forces to an invisible gravitational field, applying the conservation of energy principle to estimate these forces.

We validate that the investor impatience force effectively captures market abnormalities by examining two notably exogenous events: the GameStop short squeeze and the flash crash. In the former case, we use minute-by-minute US equity data and WallStreetBets comment sentiment estimates to illustrate how the variable enhances out-of-sample forecasting of comment sentiment. In the latter case, we demonstrate that the key feature of the flash crash, being a unified drawdown across the entire market, is also reflected in the investor impatience variable. We therefore demonstrate the macroscopic model is not only

capable of detecting market-wide events, but also adding context to events on smaller scales. This enables a potentially broad application of the model to portfolio management under unexpected events.

4.2 Related Work

Our research increments on a considerable body of literature examining the effects of world events on financial markets. Methods from multiple disciplines have been applied, including econometrics, signal processing, and physics, among others. We review approaches used for this purpose. We also discuss the application of sentiment analysis methods in financial contexts.

The versatility of econometric models in identifying relationships between time series makes unsurprising their popularity for contextualizing financial market abnormalities. An extensive body of literature applies vector autoregression (VAR) and its developments to relate global events to financial time series. Umar et al. (2022) quantifies the return and volatility connectedness between COVID-19 media coverage and segments of the non-fungible tokens market [87]. They use the same TVP-VAR approach as Antonakakis et al. (2018), which employs the Diebold and Yılmaz (2014) spillover index approach to identify volatility transmission between oil prices and the stock prices of oil and gas companies [7, 33]. Diebold and Yılmaz (2012) employs a generalized VAR approach to quantify the volatility spillover across US stock and bond, foreign exchange, and commodities markets through the global financial crisis [32]. Shahrestani and Rafei (2020) applies the Markov switching VAR model to measure the impact of oil price shocks on the Tehran Stock Exchange [78].

The literature also presents many applications of generalized autoregressive conditional heteroskedasticity (GARCH) models for detecting spillover and contextualizing market abnormalities. Dungey and Renault (2018) proposes a GARCH common features approach which is used to identify contagion during major events in the Asian currency markets,

global financial crisis, and European sovereign debt crisis [36]. The dynamic conditional correlation MGARCH model presented in Engle (2002) has been used to detect contagion via estimates of time-varying conditional correlations [38]. Among others, this method is applied to the global financial crisis in Syllignakis and Kouretas (2011) and Kim et al. (2015) to detect spillover during the global financial crisis from the US to European and emerging Asian financial markets, respectively [85, 53]. Vasileiou (2021) applies asymmetry GARCH models during the GameStop short squeeze to provide evidence of the presence of the anti-leverage effect [89].

Many authors have developed physics-inspired models of financial markets under the umbrella term *econophysics*. Such methods are frequently used in conjunction with econometric models. For a review of econophysics methods and applications, we refer the reader to the work of Chakraborti et al. (2011) [22]. The literature presents models for the detection and analysis of abnormal market events. Wavelets allow analysis of co-movements of financial data across different time scales. Ranta (2013) applies wavelets to detect contagion between major markets across decades [75]. Evidence of contagion is found during the 1987 financial crisis, the Gulf War, and the global financial crisis. Beccar et al. (2017) applies a wavelet methodology designed for geophysical data to contrast the Lehman Brothers collapse and the flash crash using minute-by-minute stock data from four companies [12]. The authors compare the former event to a natural earthquake and the latter to a human-made explosion and conclude that events of the former type are more predictable. Siddiqui et al. (2020) identifies co-movement on short time scales between major stock indices during the onset of the COVID-19 pandemic [79]. Xing et al. (2021) posits that crashes originate from changes to the underlying structure of the financial system described by the nonlinear potential function [98]. They use a GARCH model to improve forecasting of returns during market crashes. Wavelets have also been applied to cryptocurrency bubble analysis: Fruehwirt et al. (2021) identifies a structural change in relationships between cryptocurrencies towards interdependence after the 2017 Bitcoin price peak [42]. Kumar and Anandarao (2019) uses wavelet

coherence to identify volatility spillover in cryptocurrency markets [57]. Beyond only market data, Umar et al. (2021) uses X (formerly Twitter) data, the put-call ratio, and short-sale volume in a wavelet coherence approach to study the relationship between GameStop returns and sentiment during the short squeeze [88]. Log-periodic power law (LPPL) models have been used to describe market abnormalities, inspired by statistical physics and motivated by distinct groups of rational and irrational traders. Geraskin and Fantazzini (2013) provides a summary of the development and application of these approaches, beginning with the original model description from Sornette et al. (1996) [43, 81]. Wosnitza and Denz (2013) describes how LPPL structures follow the development of CDS spreads for forty banks during the 2000 financial crash [97]. Applying a LPPL model to cryptocurrency, Wheatley et al. (2019) diagnose bubbles and crashes in Bitcoin prices [96].

The phenomenon of social media, combined with advancements in natural language processing capabilities, has enabled data-driven analysis of the relationship between market action and the sentiment of market participants. This has produced a burgeoning literature on sentiment analysis applied to markets. Yang et al. (2020) used posts from the Chinese stock message board to examine the effect of investor panic on equity market crashes [102]. They used text-mining computing tools and a classification model to construct firm-level sentiment and panic indices which could predict abnormal trading and stock market crashes. Similarly, Xu et al. (2021) constructed three sentiment indices from Chinese social media, newspapers, and internet news, which have proven capable of improving forecasting of stock market returns [99]. As with forum posts, Google Search Volume Indices have been used as an indicator of investor intent. Hsieh et al. (2020) uses this data as a proxy for information demand of retail investors and to identify herding behavior of these investors [49]. Lyócsa et al. (2020) also used Google search volume to forecast global stock price variation during the COVID-19 pandemic [64]. Pedersen (2022) proposes a model of security prices in which a social network of naive, fanatical, and rational investors communicate, using the GameStop short squeeze as an example [73].

The centrality of social media during the GameStop short squeeze has resulted in multiple studies using social media data to analyze the role of sentiment in this event. Wang and Luo (2021) applies the VADER sentiment analysis package and a BERT transformer model to WallStreetBets comments in a range of classification models to predict price movements of the GameStop during the short squeeze [93, 51, 31]. Long et al. (2022) also applies the VADER package to WallStreetBets comments to quantify the relationship between sentiment and GameStop returns [63]. Mancini et al. (2022) applies the VADER package to model the dynamics of emerging consensus within WallStreetBets during the short squeeze, making a comparison between GameStop's stock price and the transition to homogeneous opinions [65].

Despite a plethora of models for detecting and forecasting abnormalities in markets, the literature lacks models approaching the problem with a macroscopic perspective. While current models capture isolated phenomena due to unique events, broader implications may be left unaddressed. The scope of recommendations along the lines of portfolio management, policy, or risk management may therefore be limited. Our model addresses this research gap by constructing a market-wide sensor for abnormalities. We conduct sentiment analysis on WallStreetBets comments in a similar vein to the aforementioned studies to demonstrate our model captures the sentiment-induced abnormality of the GameStop short squeeze. The following section details the methods with which we construct said sensor and the data used for model fitting and validation.

4.3 Methods

4.3.1 The Macroscopic Model of Equity Markets

In this research, we build upon the research of AlShelahi and Saigal, focusing on the macroscopic model of markets [5]. Our approach closely aligns with the setting outlined in their paper; we offer a brief description of this setting for context. The model uses an Eulerian

fluid-flow description of markets in which each stock's position is represented by its price. Taking $x \in \mathbb{R}_+$ to be a particular price, we construct $\rho(x,t)$ to be the density of stocks at price x and time t. Letting $N(x,t)$ denote the number of stocks in price section $[x_1, x_2]$ (with $x_1 \leq x \leq x_2$) at time t, we obtain the following definition for density:

$$N(x,t) = \int_{x_1}^{x_2} \rho(x,t)\, dx. \tag{4.1}$$

Density contextualizes the magnitude of shocks on markets: a shock to prices with greater density implies a higher impact. We may contrast density in this model with its physical interpretation. Whereas density classically refers to mass per unit volume, in this model density becomes the stocks per unit price. The velocity $v_k(t)$ of stock k at time t, is defined as

$$v_k(t) = \lim_{\Delta t \to 0} \frac{p_k(t) - p_k(t - \Delta t)}{\Delta t}, \tag{4.2}$$

where $p_k(t)$ is the price of stock k at time t. This parameter is similar to drift in classical stochastic differential equation models of stock prices. Indeed, the model may also be applied to logarithmically transformed prices. We note that our results are similar with and without such a transform and we are opting for the latter for simplicity.

The average velocity of stocks in price interval $[x_1, x_2]$ at time t can be expressed as

$$v(x,t) = \frac{1}{N(x,t)} \sum_{k\,:\,p_k(t) \in [x_1, x_2]} v_k(t), \tag{4.3}$$

as can the average squared velocity, denoted as

$$v^2(x,t) = \frac{1}{N(x,t)} \sum_{k\,:\,p_k(t) \in [x_1, x_2]} v_k^2(t). \tag{4.4}$$

We may assume that stocks are neither created nor destroyed, and thus the number of

stocks can only change from flowing across endpoints of the interval $[x_1, x_2]$. This is because stock splits, delistings, and IPOs occur rarely. We may therefore define the flux (rate of flow) of stocks at point (x, t) as $Q(x, t)$ with

$$Q(x, t) = \rho(x, t)v(x, t). \tag{4.5}$$

Flux intuitively measures the scale of an event on stock prices, increasing in both the number of stocks impacted and the rate of price changes.

The conservation of mass principle can be applied to obtain an expression for pressure as a function of flux and velocity:

$$P(x, t) = \alpha Q(x, t)v(x, t), \tag{4.6}$$

where α is a fixed parameter. This pressure is a result of the momentum with which stocks are moving in the price domain. The empirical value of α was found to be 0.3, which is carried forward to this analysis [5].

Acceleration may be defined for a stock in a similar fashion to velocity:

$$a_k(t) = \lim_{\Delta t \to 0} \frac{v_k(t) - v_k(t - \Delta t)}{\Delta t}, \tag{4.7}$$

resulting in the following expression of average acceleration:

$$a(x, t) = \frac{1}{N(x, t)} \sum_{k\,:\,p_k(t) \in [x_1, x_2]} a_k(t). \tag{4.8}$$

The following subsection extends this to derive the gravitational parameter and conservation of energy equations to estimate external forces on financial markets.

4.3.2 Formulating External Market Forces as Gravity

As detailed above, stock prices are influenced by a multitude of factors, endogenous and exogenous. We assume that external forces on stock prices due to various events act similarly to gravity in physics. Unlike gravitational force in physics, the external force on markets can operate in either direction, whether to increase or decrease stock prices. We consider two opposing masses representing positive and negative investor impatience, each with their own gravitational field. The relative size of each determines the net investor impatience.

We define the net force on stocks as the sum of internal and external forces as per Newton's second law of motion:

$$F(x,t) = m(x,t)\,a(x,t) + m(x,t)\,g(x,t), \tag{4.9}$$

where F represents the net force acting on the stocks at price x and time t, each with mass m, internal acceleration a, and external acceleration g. The mass may be expressed as the product of the stock fluid density, ρ, and volume, V:

$$m(x,t) = \rho(x,t)V(x,t). \tag{4.10}$$

Similarly, the force term may be expressed as a product of pressure and area, A:

$$F(x,t) = P(x,t)A(x,t). \tag{4.11}$$

Combining the above equations, we have

$$\alpha Q(x,t)v(x,t)A(x,t) = \rho(x,t)V(x,t)a(x,t) + \rho(x,t)V(x,t)g(x,t). \tag{4.12}$$

Taking both V and A to be unit values, and combining the formula above with the

definition of flux, we obtain

$$\alpha v^2(x,t) = a(x,t) + g(x,t). \tag{4.13}$$

Defining gravity allows for an expression of the potential energy of the market, allowing us to analyze the degree to which energy is conserved. As with momentum, the degree to which conservation is obtained may provide utility in sensing abnormal market events. We may define the potential energy of a particular stock as

$$E_k^{(P)}(t) = \rho(p_k(t),t)g(p_k(t),t)p_k(t) \tag{4.14}$$

where p_k is the price of stock k. The average potential energy in a particular price interval $[x_1, x_2]$ can therefore be defined as

$$E^{(P)}(x,t) = \frac{1}{N(x,t)} \sum_{k\,:\,p_k(t)\in[x_1,x_2]} E_k^{(P)}(t). \tag{4.15}$$

We likewise define the kinetic energy of the stock, assumed to have unit mass as

$$E_k^{(K)}(t) = \frac{1}{2} v_k^2(t). \tag{4.16}$$

Consequently, we can quantify the kinetic energy of the price interval as

$$E^{(K)}(x,t) = \frac{1}{N(x,t)} \sum_{k\,:\,p_k(t)\in[x_1,x_2]} E_k^{(K)}(t). \tag{4.17}$$

Total energy at this price interval may therefore be specified as

$$E(x,t) = E^{(K)}(x,t) + E^{(P)}(x,t). \tag{4.18}$$

Applying the Euler energy conservation equation in one dimension, and assuming that

energy, density, and velocity are all differentiable, we thus have

$$\frac{\partial E(x,t)}{\partial t} + \frac{\partial}{\partial x}((E(x,t) + P(x,t))v(x,t)) = 0 \quad \forall x, t > 0. \tag{4.19}$$

The complete model consists of three stochastic partial differential equations representing the conservation of mass, momentum, and energy principles. The following subsection details how we fit the model to data from the short squeeze and flash crash.

4.3.2.1 Data

The Data For analysis of the Gamestop short squeeze, we use minute-by-minute price data collected from *Yahoo! Finance* for 4494 stocks listed on US exchanges. Each stock was included only if data were available for at least 75% of the period from January 21 through January 29, 2021, resulting in an effective sample size of 1690 stocks. When data for a particular minute was unavailable, we imputed the price from the most recently available price. Due to the sparsity of stocks with prices over $100, we analyze only those below this threshold. We obtain similar data for the flash crash and are left with an effective sample size of 2853 stocks following the same data-cleaning procedure.

We use the Pushshift Reddit dataset to obtain all available WallStreetBets comments during market hours from January 21 through 29 [11]. We exclude comments that have been deleted or removed, contain URLs, or were authored by the AutoModerator, an automated moderation tool. We further pre-process the comments by substituting usernames with '@user', cleaning the text of escape sequences, and converting HTML entities to their respective characters. The dataset contains approximately 1.8 million comments made during market hours.

Discretization and Fitting We apply the following discretizing and fitting approach for parameter estimation. A resolution of $1 and 1 minute is applied. Consistent with [5], a linear regression model is fitted relating observed acceleration to squared velocity, estimating

gravity for a price/time combination of these variables. We have

$$a(x,t) = -g(x,t) + \beta\alpha v^2(x,t) + \sigma(x,t)\epsilon, \tag{4.20}$$

where $\epsilon \sim \mathcal{N}(0,1)$ is an error term. To estimate $g(x,t)$, the regression uses calculated values for a and v^2 in the 3×3 grid of prices $\{(x_i, t_j) : i \in \{x-2, x-1, x\}, j \in \{t-2, t-1, t\}\}$.

We define a series $g_k(t)$ to be the external acceleration from the perspective of stock k:

$$g_k(t) = g(x,t), \quad p_k(t) \in S_x. \tag{4.21}$$

where S_x is the discretized price section containing x. This highlights the applicability of the macroscopic model to phenomena on lower scales. Individual stocks' price changes can be contextualized within the broader macroscopic environment.

Parameterizing the Energy Conservation Equation Although the conservation equation (4.19) right-hand side is zero at all points, a forcing term is required due to the discretization error and uncertainty which disturb this equation. We therefore propose a stochastic forcing term as follows:

$$\frac{\partial E(x,t)}{\partial t} + \frac{\partial}{\partial x}\left((E(x,t) + P(x,t))v(x,t)\right) = z(x,t) \tag{4.22}$$

$$z(x,t) = l(x,t) + \eta(x,t)E(x,t) + \sigma_3(x,t)\frac{dW_3(x,t)}{dxdt} \tag{4.23}$$

Here, $l(x,t)$ is a deterministic function representing the mean inflow or outflow of the right-hand side of (4.22). The $\eta(x,t)$ term is likewise a deterministic function for the rate of reversion to the mean of the right-hand side. The $W_3(x,t)$ term is a Brownian sheet, a Gaussian stochastic process with mean 0 and covariance $\mathbb{E}\left(W_3(x_1,t_1)W_3(x_2,t_2)\right) = \min(x_1,x_2)\cdot\min(t_1,t_2)$ [92]. The $\sigma_3(x,t)$ term represents the volatility of the energy conservation process.

4.3.2.2 Estimating Sentiment from Reddit Data

Introduction We apply two techniques to estimate the sentiment of WallStreetBets comments. The first is VADER, a lexicon- and rule-based sentiment analysis tool designed specifically for social media text [51]. We also apply a pre-tuned version of the RoBERTa transformer-based machine learning model [51, 62, 50]. We use the Twitter-RoBERTa-base-sentiment model from CardiffNLP to estimate comment sentiment [50]. The RoBERTa-base model is pre-trained on the English Wikipedia and BookCorpus datasets [62]. The Twitter-RoBERTa-base model is then tuned on a set of approximately 58 million tweets [9]. Tasks this model is trained for include sentiment analysis, irony detection, and hate speech detection, among others.

Although a RoBERTa model tuned on WallStreetBets comments would improve performance, there were none available at the time of our analysis. We consider the RoBERTa-base model tuned on X (formerly Twitter) data appropriate for this analysis as both Reddit and X are social media websites and are thus likely to share commonalities. The similarity between the X tuning data and our Reddit data is corroborated by our observation that, despite the option of long-form comments on Reddit, approximately 89% of comments in our dataset are within X's original character limit of 140 characters. However, we acknowledge the limitations of this approach. Despite both being social media websites, X and WallStreetBets have unique cultures. WallStreetBets comments are often ironic, esoteric, and loaded with forum-specific references and inside jokes. The model may therefore misclassify some comments from WallStreetBets. Further analysis may be required to both evaluate the model's accuracy on WallStreetBets comments and, if required, tune a model for use on WallStreetBets comments. The following subsections detail how comment sentiment is estimated and the forecasting procedure.

Sentiment Estimation using VADER The VADER sentiment package returns a normalized, weighted composite score for each comment's sentiment between 1 (most positive)

and -1 (most negative). We denote the sentiment score of comment i as $\bar{s}_i$ and the time the comment was posted by $\hat{t}_i$. We construct a minute-by-minute estimate of the overall sentiment of the WallStreetBets forum by calculating the mean sentiment of all comments posted in each minute. We denote the average sentiment at minute t, given by $S_{\text{VADER}}(t)$, to be

$$S_{\text{VADER}}(t) = \frac{1}{|\{i : \hat{t}_i \in [t, t+1)\}|} \sum_{i:\hat{t}_i \in [t,t+1)} \bar{s}_i. \tag{4.24}$$

Sentiment Estimation using the Twitter-RoBERTa-base Sentiment Model The Twitter-RoBERTa-base model for sentiment analysis outputs values corresponding to the labels negative, neutral, and positive. Denoting $\bar{n}_i$ and $\bar{p}_i$ to be the softmax transformations of the negative and positive output values for comment i, respectively, we estimate the average sentiment at minute t, denoted by $S_{\text{roBERTa}}(t)$, to be

$$S_{\text{roBERTa}}(t) = \frac{1}{|\{i : \hat{t}_i \in [t, t+1)\}|} \sum_{i:\hat{t}_i \in [t,t+1)} (\bar{p}_i - \bar{n}_i). \tag{4.25}$$

Forecasting Comment Sentiment We use the Bayesian time series forecasting and inference package Orbit to predict the WallStreetBets comment sentiment estimations [70]. To construct a smaller set of regressors from the investor impatience field, we take the average investor impatience across price subsets. We choose the price interval of \$2 to \$30, as these price sections generally contain at least 10 to 15 stocks per dollar. We construct four average investor impatience series, denoted by $\bar{g}_i(t), i \in \{1, 2, 3, 4\}$ and defined as

$$\bar{g}_i(t) = \frac{1}{7} \sum_{j=1}^{7} g(2 + 7(i-1) + j, t), \quad i \in \{1, 2, 3, 4\}. \tag{4.26}$$

On each trading day from January 21 through 29, we apply Orbit's BackTester method on the Damped Local Trend (DLT) model using an expanding window with a linear global trend, minimum training window length of 120 minutes, a forecast length of 30 minutes, and

an increment length of 30 minutes. The number of samples for each model fit is 500. We use the symmetric mean absolute percentage error (SMAPE) as the error metric, since Orbit's models outperform in terms of this metric compared to other candidate time series models [70]. For each sentiment metric and each trading day, two models are fitted: one DLT model for the sentiment metric, and one using the investor impatience series as regressors for the sentiment metric.

4.4 Results and Discussion

4.4.1 Introduction

We first qualitatively analyze the investor impatience force across the GameStop short squeeze and flash crash. A macroscopic perspective is applied to notable days during these events to demonstrate the utility of the model's scale and comment on the model's potential use cases. This is followed by contextualizing individual stocks' price action during the short squeeze to show applicability of the model to smaller scales. We also demonstrate the potential for the model to serve as a sensor of market abnormalities via the conservation of energy equation parameters during the flash crash. Lastly, we present results from sentiment forecasting to quantitatively confirm the model is capable of detecting market abnormalities.

4.4.2 Macroscopic Investor Impatience

Figure 4.1 below shows the investor impatience field for January 25 during the short squeeze. Also plotted are the prices of three meme stocks of interest: AMC Entertainment Holdings Inc. ($AMC), Bed Bath & Beyond ($BBBY), and Eastman Kodak Company ($KODK). We note that the color scale limits have been adjusted for visual aid due to the presence of outliers. Meme stocks saw large price increases before sustaining sharp declines beginning around 10:45am. The investor impatience variable captures the sharp mid-morning drawdown of the meme stocks. The vertical striations of color display unity of price action

across multiple price sections, demonstrating the model's ability to capture market-wide phenomena.

The equivalent graph for during the flash crash is shown in Figure 4.2 below, with the crash onset highlighted by the green dashed line. Similar color patterns are observed. Notably, a presence of wave-like negative investor impatience striations are observed prior to the onset of the crash. This corroborates the findings of [5], which indicates the model variables may be used as sensors of abnormal market activity.

Figure 4.3 below shows the investor impatience force during January 28, as well as the aforementioned stocks' prices. Once again, we see the presence of negative investor impatience values throughout the drawdown phases.

These figures indicate potential applications of this model to hedging and regulation. The unification of investor impatience across price sections appears to occur more strongly during drawdowns than rallies. This supports the empirical observation that correlations between equities are much greater for downside moves than upside moves [6]. One may therefore use historical estimates of investor impatience to inform portfolio construction. Furthermore, the ability of the model to act as a sensor for market abnormalities may be used by regulators within a broader crash detection framework to prophylactically restrict a crashing market.

However, it is yet to be seen whether broader and persistent external influences on markets can be captured by investor impatience. Economic and political contexts patently influence market prices on time scales beyond those analyzed here, warranting further investigation. Attribution of external forces, too, should be studied to determine the efficacy of the model in capturing varied market phenomena.

We explore the meme stocks' positions within the investor impatience field in the following subsection.

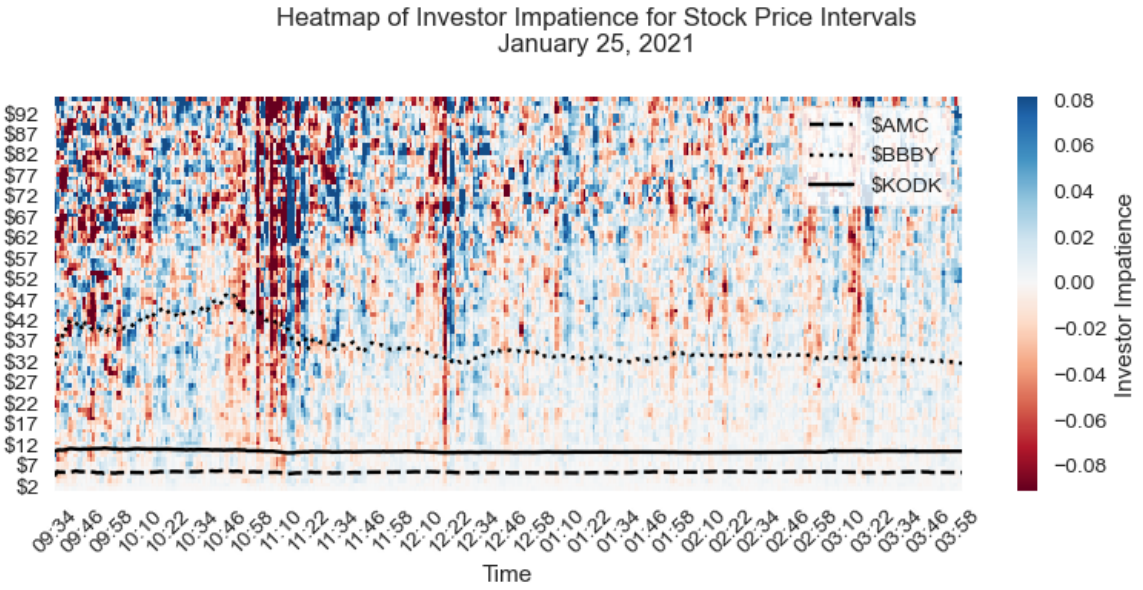

Figure 4.1: Investor Impatience Heatmap and Prices of $AMC, $BBBY, and $KODK for January 25, 2021

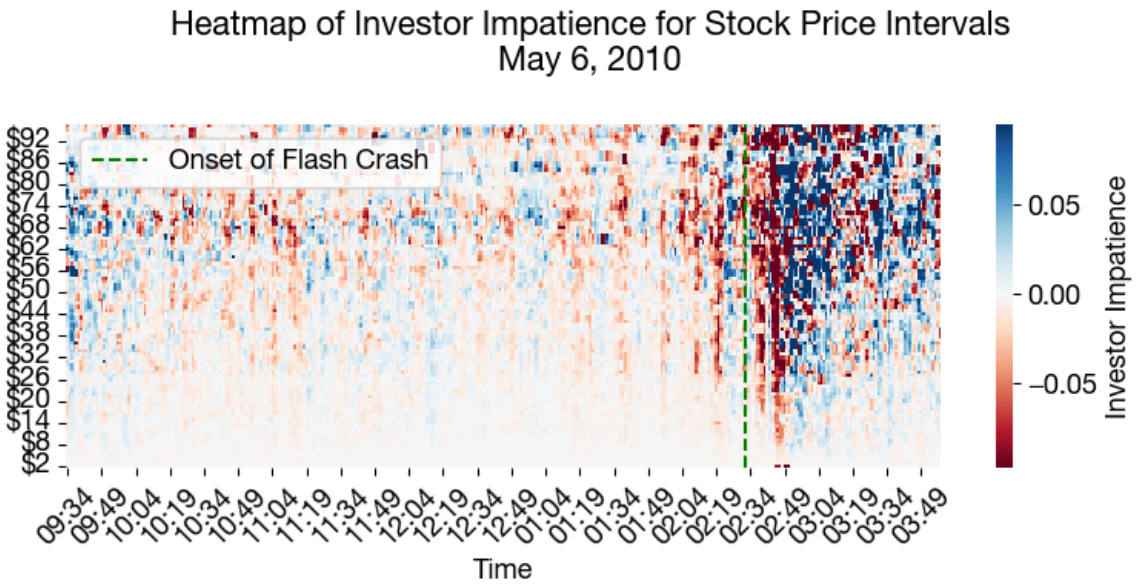

Figure 4.2: Investor Impatience for All Stocks during the Flash Crash

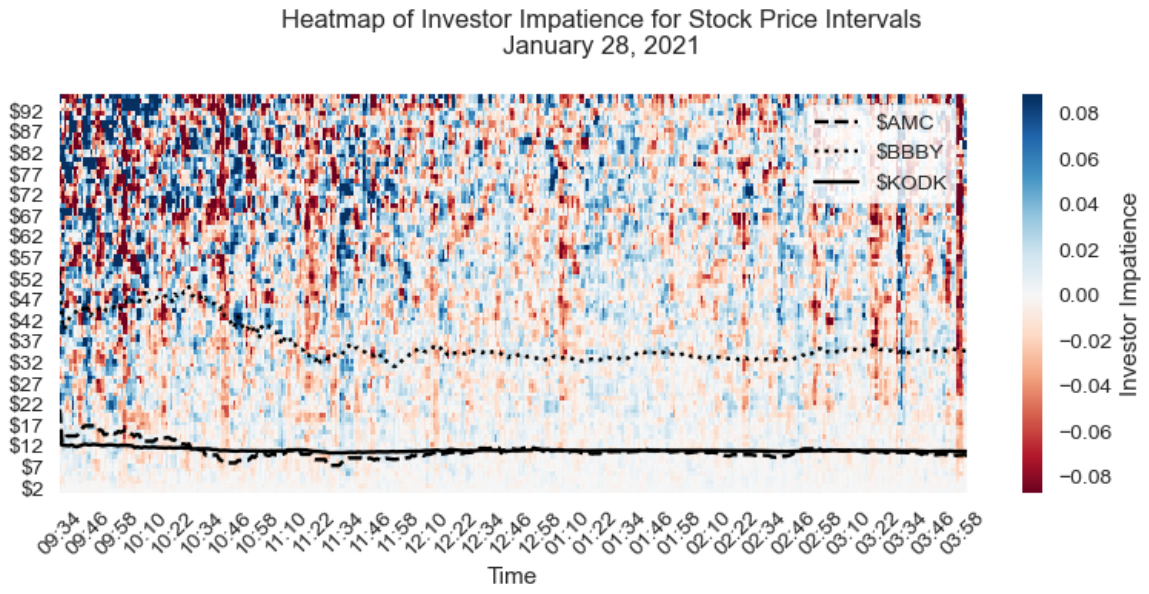

Figure 4.3: Investor Impatience Heatmap and Prices of $AMC, $BBBY, and $KODK for January 28, 2021

4.4.3 Microscopic Investor Impatience

We now examine the application of the model to individual stocks. Figures 4.4 through 4.6 below show the price, traded volume, and investor impatience force (i.e. $g_k(t)$) for Bed Bath & Beyond, AMC Entertainment Holdings Inc., and Eastman Kodak Company on January 28. Also plotted is the 10-minute rolling average of investor impatience, for clarity. We note that GameStop's price was generally above $100 and is thus excluded. For all three stocks, considerable negative investor impatience is observed during the morning drawdown.

Figure 4.7 below displays the investor impatience and conservation of energy equation parameters during the flash crash. Again, the rolling 10-minute average of investor impatience is plotted. There is not only negative investor impatience prior to the crash, but also a noticeable increase in the energy conservation volatility term. These indicate this system may be used for detecting abnormalities in markets.

These figures also reveal how one may contextualize the price action of individual stocks within the investor impatience field to "decompose" a stock's price action into components

of market-wide external force and idiosyncratic action. Such an approach may be applicable to hedging. A stock with positive acceleration during a negative investor impatience environment (and vice versa) may have implications for portfolio diversification, particularly in market phenomena that impact a particular subset of traded assets.

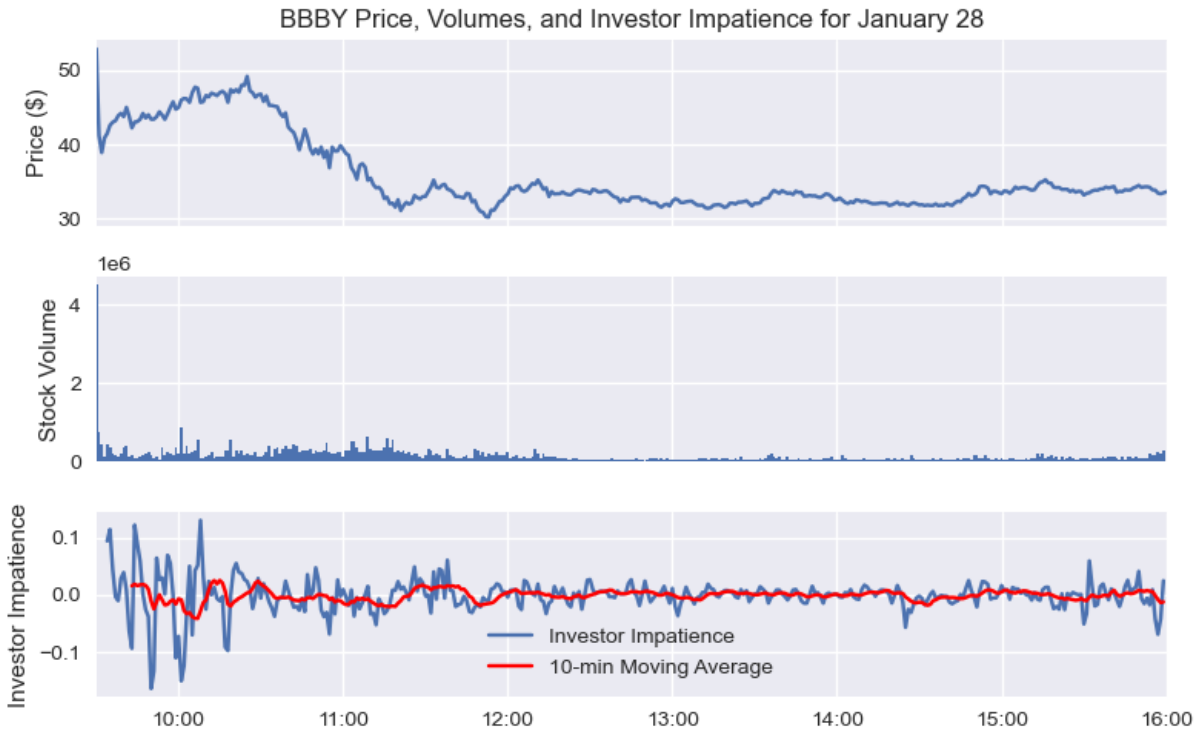

Figure 4.4: Bed Bath & Beyond ($BBBY) Price, Volume, and Investor Impatience for January 28, 2021

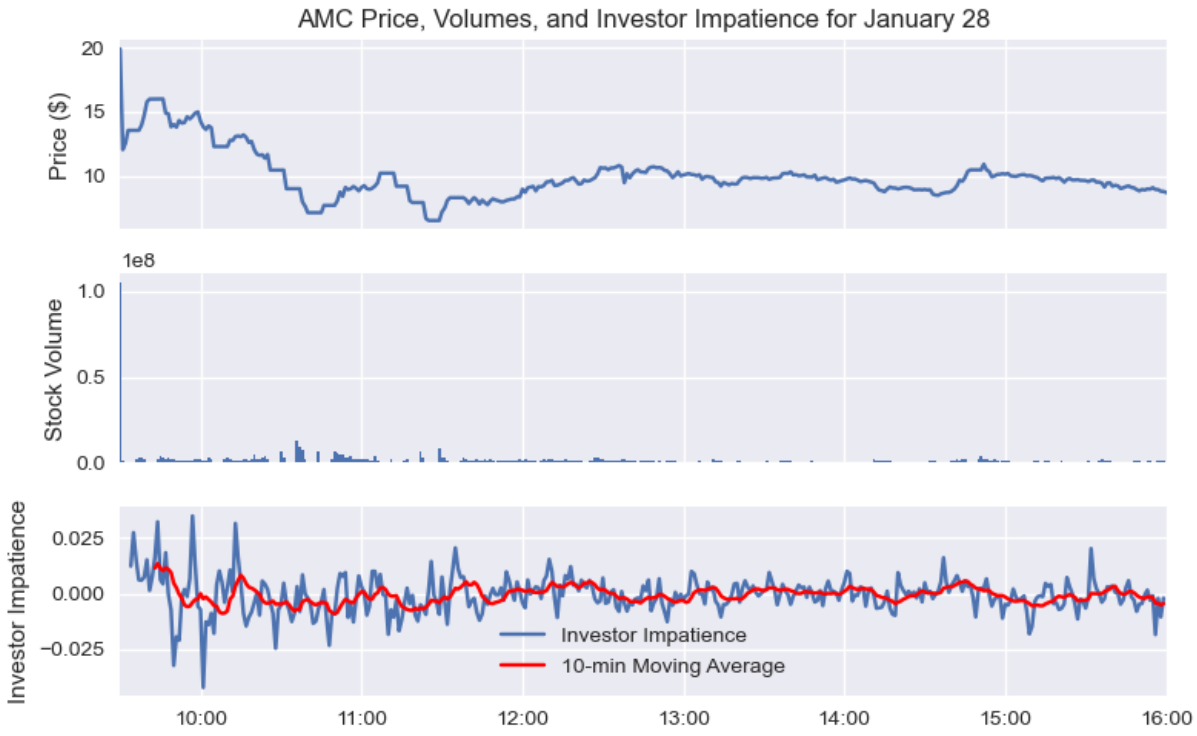

Figure 4.5: AMC Entertainment Holdings, Inc. ($AMC) Price, Volume, and Investor Impatience for January 28, 2021

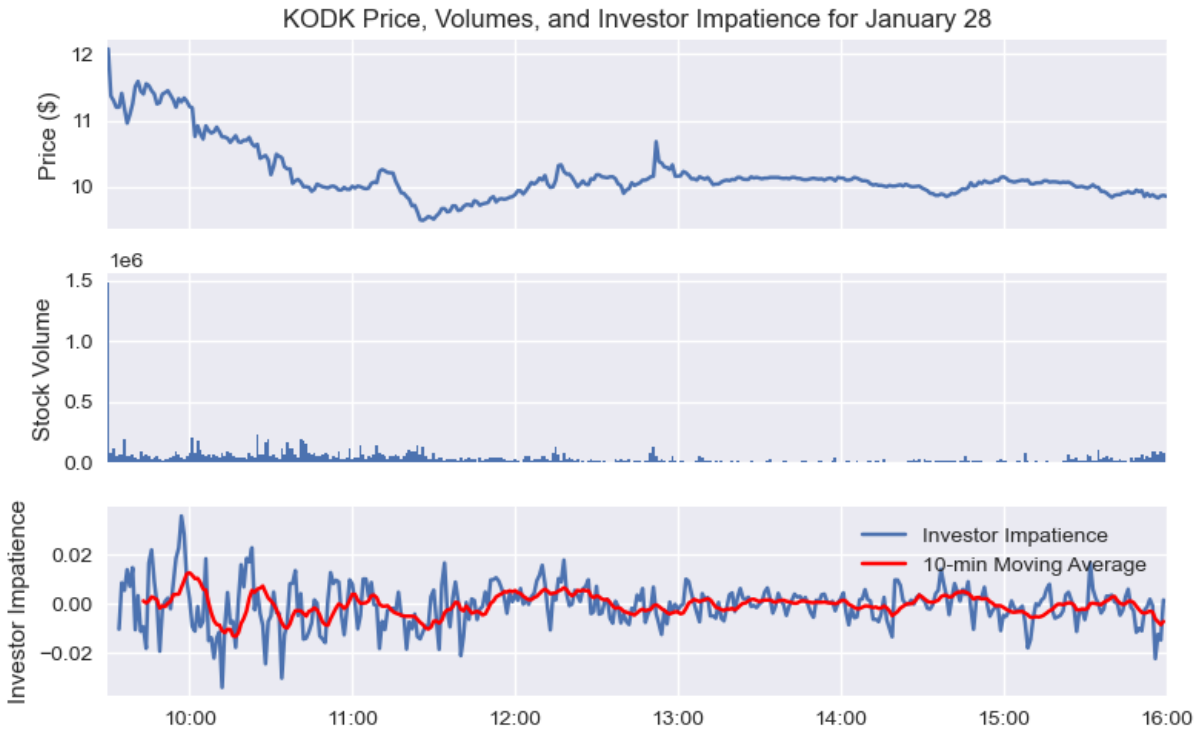

Figure 4.6: Eastman Kodak Company ($KODK) Price, Volume, and Investor Impatience for January 28, 2021

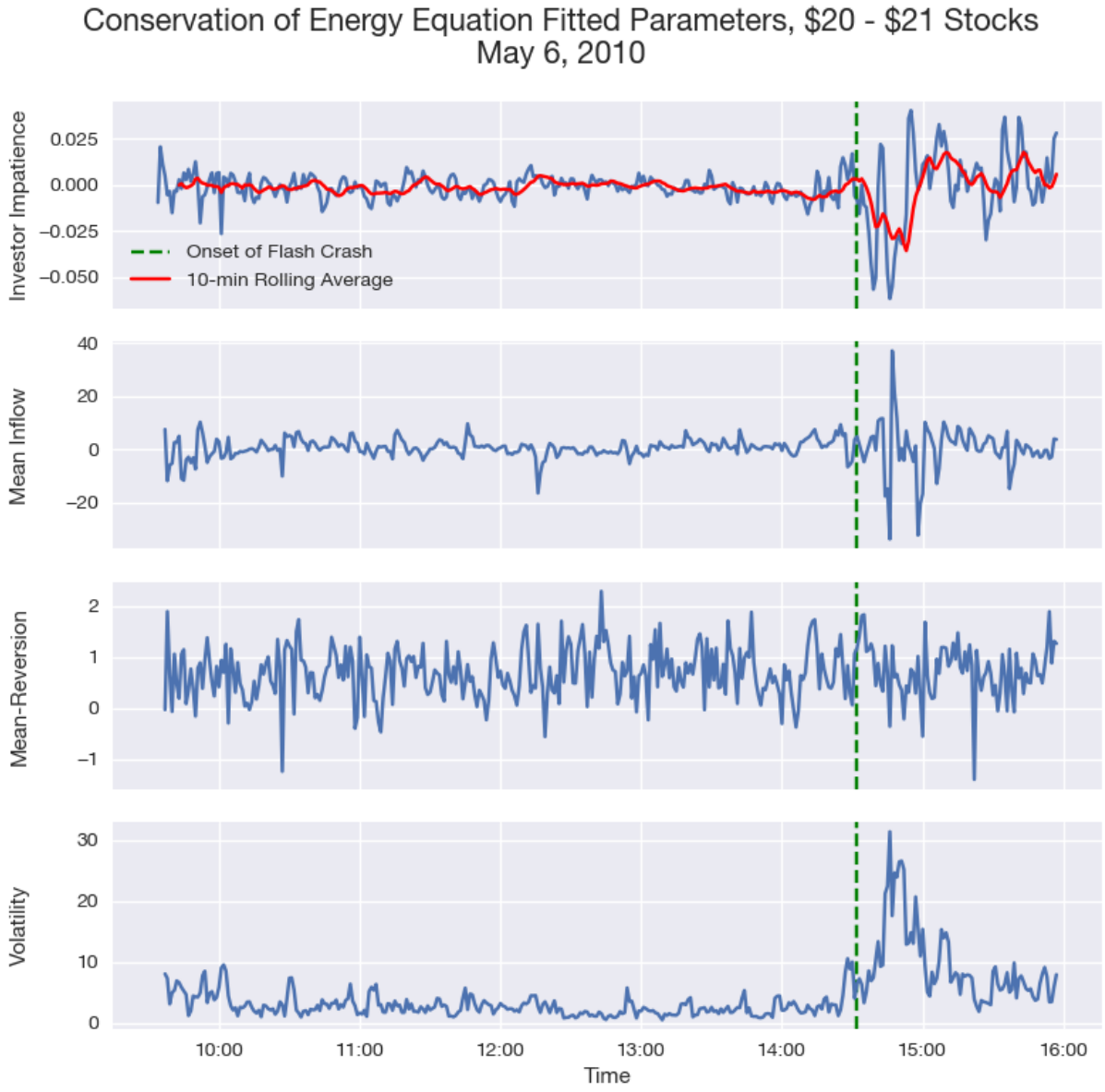

Figure 4.7: Investor Impatience and Conservation of Energy Parameters for $20 Stocks during the Flash Crash

4.4.4 Sentiment Forecasting

Table 4.1 below shows the out-of-sample SMAPE metric percentage improvement using the investor impatience regressors for each trading date using both sentiment models. Inclusion of the investor impatience regressors decreases the prediction error in five of seven days for

both estimates of sentiment, although performance is superior for the Twitter-RoBERTa-base estimates. These results indicate the model parameters capture external phenomena that may have market impact, such as the role of investor sentiment in the short squeeze.

	Sentiment Estimation Method	
Trading Day	**VADER**	**Twitter-RoBERTa-base**
January 21, 2021	-1.36%	-1.45%
January 22, 2021	-1.85%	2.38%
January 25, 2021	-5.49%	-4.18%
January 26, 2021	-5.52%	-0.37%
January 27, 2021	2.61%	-1.60%
January 28, 2021	8.99%	1.42%
January 29, 2021	-1.30%	-6.06%
Average	**-0.56%**	**-1.41%**

Table 4.1: SMAPE Prediction Error Change from Including Investor Impatience Regressors during the GameStop Short Squeeze

4.5 Conclusions and Future Work

We have estimated external forces on equity markets under a physics-based macroscopic model. We fitted the investor impatience force, which is conceptually similar to the gravitational force, and the conservation of energy equation, analyzing them alongside a selection of stocks that exhibited abnormal behavior during the GameStop short squeeze. Our results indicate this model captures a degree of investor sentiment during this event and can be used as a sensor of abnormal market activity. The latter point is supported by our qualitative analysis of investor impatience and the energy conservation equation during the flash crash. Furthermore, the comparable results across contrasting market events indicates the ability of the model to respond to a potential multitude of external influences. Multiple future avenues

of research are presented by these findings. Naturally, future work may evaluate the investor impatience variable and energy equation as predictive tools of market crashes. Similar analyses may be conducted on fixed income or cryptocurrency markets. Further analysis on the potential hedging capabilities of various portfolios from an investor impatience perspective is also warranted.

CHAPTER 5

Conclusion

The challenge of making the right long-term investments in financial markets stems from the complexity of not only the investor's desires and financial situation, but also the markets themselves. This dissertation reduces the complexity from both such sources. Other GBWM models have lacked key elements of realistic lifetime goal-setting, including retirement considerations and a connection between income and the market at large. The first part of this dissertation introduces a model which addresses such considerations and beyond. We solve this model for a hypothetical investor and demonstrate the quantitative and qualitative difference in their optimal strategy as their situation changes. The second part of this dissertation presents a reinforcement learning approach to GBWM that addresses recession risk and uncertain goal times. The model demonstrates sensitivity to both of these factors, providing intuitive recommendations for investment and goal contributions. The final part of this dissertation works to measure external forces on financial markets by relating such forces to the physical phenomenon of gravity. Within a fluid dynamics model, we measure these forces using the conservation of energy principle, represented as a stochastic PDE. We validate the model with minute-by-minute stock and forum comment data during the GameStop short squeeze, as well as a separate study on the flash crash.

While this research has made progress towards the understanding of, and decision-making within, financial markets, there are multiple natural extensions. Refinements to the GBWM model in Chapter 2 can be made in a few areas. Nuances in tax codes and retirement

account withdrawals are not yet addressed, nor the spending within the post-retirement phase of life. In addition to these, incorporating market regimes such as recessions would be useful model improvements. For additional realism, the assumption of geometric Brownian motions underlying all processes could be relaxed in future research. The model presented in Chapter 3 can be expanded to a more comprehensive framework, incorporating elements from Chapter 2 and other complexities. For Chapter 4, extensions may include studies of other markets, such as those of fixed-income or cryptocurrency. Synthesis of the ideas presented in this dissertation may include exploring optimal decision-making in the presence of external forces on financial markets. The nonlinearities expected under exogenous market impacts may provide insights into investment strategy for risk-averse investors.

APPENDIX A

Finite Difference Approximation of the Value Function

A.1 Introduction

Prior to applying the shape-preserving Chebyshev polynomials, a finite difference approach was tried, using an Euler scheme to approximate the value function. For completeness, this appendix details the method used despite its application being unsuccessful.

Issues arose in maintaining concavity of the value function estimates in successive time steps. We expect this is due to a combination of numerical error and the value function shape. For many states in which maximal future utility is all but guaranteed, such as high X_P and X_R values for moderate income levels, the portfolio-dependent derivatives approach zero. The $\frac{\partial^2 F}{\partial X_P^2}$ derivative estimate becomes positive at multiple points, rendering the problem nonconcave. This led to abnormal portfolio recommendations. Attempts to smooth the value function or enforce concavity in the diagonals of the value function Hessian (either within or after the policy iteration) did not result in sensible portfolios.

The following subsections detail the finite difference discretization and policy iteration procedure.

A.1.1 Discretization

The state space is discretized into a uniform mesh. The set of states in each program, $\mathbf{X}_k$, $k = 1, 2, ..., K$ can be defined with:

$$\mathbf{t} = [0, \Delta_t, 2\Delta_t, ..., N_t\Delta_t] \tag{A.1}$$

$$\mathbf{P} = [0, \Delta_P, 2\Delta_P, ..., N_P\Delta_P]^T \tag{A.2}$$

$$\mathbf{R} = [0, \Delta_R, 2\Delta_R, ..., N_R\Delta_R]^T \tag{A.3}$$

$$\mathbf{I} = [0, \Delta_I, 2\Delta_I, ..., N_I\Delta_I]^T \tag{A.4}$$

$$\mathbf{X}_k = \{(t, X_P, X_R, X_I) : t_{k-1} \leq t < t_k, t \in \mathbf{t}, X_P \in \mathbf{P}, X_R \in \mathbf{R}, X_I \in \mathbf{I}\}, \quad k = 1, 2, ..., K \tag{A.5}$$

$$\tag{A.6}$$

where: N_i, $i = t, X_P, X_R, X_I$ represents the number of discretized points for process i; $\Delta_i > 0$; $t_k \in \mathbf{t}$ for $k = 1, 2, ..., K-1$; and $(N_t + 1)\Delta_t = t_K$.

For notational clarity, we include each component of the financial state X as a separate input to the value functions: $F_k(t, X_P, X_R, X_I)$. Derivatives of the value functions are

approximated using an Euler scheme as follows, for $(t, X_P, X_R, X_I) \in \mathbf{X}_k$, $k = 1, 2, ..., K$.

$$\frac{\partial F_k(t, X_P, X_R, X_I)}{\partial t} \approx \frac{F_k(t + \Delta_t, X_P, X_R, X_I) - F_k(t, X_P, X_R, X_I)}{\Delta_t} \tag{A.7}$$

$$\frac{\partial F_k(t, X_P, X_R, X_I)}{\partial X_P} \approx \frac{F_k(t, X_P + \Delta_P, X_R, X_I) - F_k(t, X_P, X_R, X_I)}{\Delta_P} \tag{A.8}$$

$$\frac{\partial F_k(t, X_P, X_R, X_I)}{\partial X_R} \approx \frac{F_k(t, X_P, X_R + \Delta_R, X_I) - F_k(t, X_P, X_R, X_I)}{\Delta_R} \tag{A.9}$$

$$\frac{\partial F_k(t, X_P, X_R, X_I)}{\partial X_I} \approx \frac{F_k(t, X_P, X_R, X_I + \Delta_I) - F_k(t, X_P, X_R, X_I)}{\Delta_I} \tag{A.10}$$

$$\frac{\partial^2 F_k(t, X_P, X_R, X_I)}{\partial X_P^2} \approx \frac{F_k(t, X_P - \Delta_P, X_R, X_I) - 2F_k(t, X_P, X_R, X_I) + F_k(t, X_P + \Delta_P, X_R, X_I)}{\Delta_P^2} \tag{A.11}$$

$$\frac{\partial^2 F_k(t, X_P, X_R, X_I)}{\partial X_R^2} \approx \frac{F_k(t, X_P, X_R - \Delta_R, X_I) - 2F_k(t, X_P, X_R, X_I) + F_k(t, X_P, X_R + \Delta_R, X_I)}{\Delta_R^2} \tag{A.12}$$

$$\frac{\partial^2 F_k(t, X_P, X_R, X_I)}{\partial X_I^2} \approx \frac{F_k(t, X_P, X_R, X_I - \Delta_I) - 2F_k(t, X_P, X_R, X_I) + F_k(t, X_P, X_R, X_I + \Delta_I)}{\Delta_I^2} \tag{A.13}$$

$$\frac{\partial^2 F}{\partial X_P \partial X_R} \approx \frac{1}{4\Delta_P \Delta_R} \left(F_k(t, X_P + \Delta_P, X_R + \Delta_R, X_I) - F_k(t, X_P - \Delta_P, X_R + \Delta_R, X_I)\right. \tag{A.14}$$

$$\left. - F_k(t, X_P + \Delta_P, X_R - \Delta_R, X_I) + F_k(t, X_P - \Delta_P, X_R - \Delta_R, X_I)\right) \tag{A.15}$$

$$\frac{\partial^2 F}{\partial X_P \partial X_I} \approx \frac{1}{4\Delta_P \Delta_I} (F_k(t, X_P + \Delta_P, X_R, X_I + \Delta_I) - F_k(t, X_P - \Delta_P, X_R, X_I + \Delta_I) \tag{A.16}$$

$$- F_k(t, X_P + \Delta_P, X_R, X_I - \Delta_I) + F_k(t, X_P - \Delta_P, X_R, X_I - \Delta_I)) \tag{A.17}$$

$$\frac{\partial^2 F}{\partial X_R \partial X_I} \approx \frac{1}{4\Delta_R \Delta_I} (F_k(t, X_P, X_R + \Delta_R, X_I + \Delta_I) - F_k(t, X_P, X_R - \Delta_R, X_I + \Delta_I) \tag{A.18}$$

$$- F_k(t, X_P, X_R + \Delta_R, X_I - \Delta_I) + F_k(t, X_P, X_R - \Delta_R, X_I - \Delta_I)) \tag{A.19}$$

where we have the following boundary conditions to maintain nonpositive second derivative approximations:

$$F_k(t,(N_P+1)\Delta_P,X_R,X_I) = F_k(t,N_P\Delta_P,X_R,X_I) \qquad t\in\mathbf{t},X_R\in\mathbf{R},X_I\in\mathbf{I} \tag{A.20}$$

$$F_k(t,X_P,(N_R+1)\Delta_R,X_I) = F_k(t,X_P,N_R\Delta_R,X_I) \qquad t\in\mathbf{t},X_P\in\mathbf{P},X_I\in\mathbf{I} \tag{A.21}$$

$$F_k(t,X_P,X_R,(N_I+1)\Delta_I) = F_k(t,X_P,X_R,N_I\Delta_I) \qquad t\in\mathbf{t},X_P\in\mathbf{P},X_R\in\mathbf{R} \tag{A.22}$$

$$F_k(t,-\Delta_P,X_R,X_I) = -F_k(t,\Delta_P,X_R,X_I) \qquad t\in\mathbf{t},X_R\in\mathbf{R},X_I\in\mathbf{I} \tag{A.23}$$

$$F_k(t,X_P,-\Delta_R,X_I) = -F_k(t,X_P,\Delta_R,X_I) \qquad t\in\mathbf{t},X_P\in\mathbf{P},X_I\in\mathbf{I} \tag{A.24}$$

$$F_k(t,X_P,X_R,-\Delta_I) = -F_k(t,X_P,X_R,\Delta_I) \qquad t\in\mathbf{t},X_P\in\mathbf{P},X_R\in\mathbf{R} \tag{A.25}$$

A.1.2 Policy Iteration

Algorithm 1 below details the policy iteration procedure applied to this finite difference scheme.

Algorithm 1 Policy Iteration

Require: $\epsilon_1, \epsilon_2, N > 0$
  $k \leftarrow K$
  while $k \geq 1$ **do**
    Calculate $G_k^\star(X_P, X_R, X_I)$ for all X_P, X_R, X_I via enumeration (taking $G_k^\star = \max_{G_k \in \mathbf{G}_k^\star} G_k$ where $\mathbf{G}_k^\star$ denotes the set of maximizers if non-unique)
    $F_k(t_k, \cdot, \cdot, \cdot) \leftarrow \Phi_k(\cdot, \cdot, \cdot)$
    $t \leftarrow t_k - \Delta_t$
    while $t \geq t_{k-1}$ **do**
      $\delta_1 \leftarrow \infty$ ▷ Norm of the HJB residual
      $F_k(t, \cdot, \cdot, \cdot) \leftarrow F_k(t + \Delta_t, \cdot, \cdot, \cdot)$
      $(\pi(t), \lambda_P(t), \lambda_R(t), \lambda_C(t)) \leftarrow (\pi(t + \Delta_t), \lambda_P(t + \Delta_t), \lambda_R(t + \Delta_t), \lambda_C(t + \Delta_t))$
      while $\delta_1 > \epsilon_1$ **do**
        $\delta_2 \leftarrow \infty$ ▷ Norm of value function change
        $n \leftarrow 0$
        $F_k(t, \cdot, \cdot, \cdot) \leftarrow F_k(t + \Delta_t, \cdot, \cdot, \cdot)$
        while $n < N$ and $\delta_2 > \epsilon_2$ **do**
          $F_{\text{new}}(t, \cdot, \cdot, \cdot) \leftarrow F_{\text{iter}}(F_k(t+\Delta_t, X_P, X_R, X_I), F_k(t, X_P, X_R, X_I), \tilde{u}_k; \pi, \lambda_P, \lambda_R, \lambda_C)$ ▷ per (A.26)
          $\delta_2 \leftarrow \|F_{\text{new}}(t, \cdot, \cdot, \cdot) - F(t, \cdot, \cdot, \cdot)\|_2$
          $F_k(t, \cdot, \cdot, \cdot) \leftarrow F_{\text{new}}(t, \cdot, \cdot, \cdot)$
          $n \leftarrow n + 1$
        end while
        $\mathbf{\Delta} \leftarrow$ Numerical derivatives of $F(t, \cdot, \cdot, \cdot)$ ▷ per (A.7) to (A.19)
        Optimize $\pi, \lambda_P, \lambda_R, \lambda_C$ ▷ per (A.27)
        $\delta_1 \leftarrow \|\text{HJBResidual}(F(t, \cdot, \cdot, \cdot), \mathbf{\Delta}, (\pi(t), \lambda_P(t), \lambda_R(t), \lambda_C(t)), \tilde{u}_k)\|_2$
      end while
      $t \leftarrow t - \Delta_t$
    end while
    $k \leftarrow k - 1$
  end while

We define the function F_{iter} as follows.

$$
\begin{aligned}
&F_{\text{iter}}(F(t+\Delta_t, X_P, X_R, X_I), F(t, X_P, X_R, X_I), \tilde{u}_k; \pi, \lambda_P, \lambda_R, \lambda_C) \\
&\quad = \frac{\alpha(F(t+\Delta_t, X_P, X_R, X_I), F(t, X_P, X_R, X_I), \tilde{u}_k; \pi, \lambda_P, \lambda_R, \lambda_C)}{\beta(X_P, X_R, X_I; \pi, \lambda_P, \lambda_R, \lambda_C)} \\
&\quad \beta(X_P, X_R, X_I; \pi, \lambda_P, \lambda_R, \lambda_C) \\
&\qquad = r + \frac{1}{\Delta_t} + \frac{1}{\Delta_P}((r(1-\pi^T \mathbb{1}_N) + \pi^T \mu)X_P + (1-\nu_I)\lambda_P X_I) \\
&\qquad + \frac{1}{\Delta_R}(\mu_R X_R + (1+\kappa)\lambda_R X_I) + \frac{1}{\Delta_I}(\mu_I X_I) + \frac{1}{\Delta_P^2} X_P^2 \pi^T C_{PP} \pi \\
&\qquad + \frac{1}{\Delta_R^2} X_R^2 C_{RR} + \frac{1}{\Delta_I^2} X_I^2 C_{II} \\
&\quad \alpha(F(t+\Delta_t, X_P, X_R, X_I), F(t, X_P, X_R, X_I), \tilde{u}_k; \pi, \lambda_P, \lambda_R, \lambda_C) \\
&\qquad = \frac{F(t+\Delta_t, X_P, X_R, X_I)}{\Delta_t} + \tilde{u}_k((1-\nu_I)\lambda_C X_I) \\
&\qquad + \frac{F(t, X_P+\Delta_P, X_R, X_I)}{\Delta_P}\left(\left(r(1-\pi^T \mathbb{1}_N) + \pi^T \mu\right) X_P + (1-\nu_I)\lambda_P X_I\right) \\
&\qquad + \frac{F(t, X_P, X_R+\Delta_R, X_I)}{\Delta_R}(\mu_R X_R + (1+\kappa)\lambda_R X_I) + \frac{F(t, X_P, X_R, X_I+\Delta_I)}{\Delta_I}\mu_I X_I \\
&\qquad + \frac{F(t, X_P+\Delta_P, X_R, X_I) + F(t, X_P-\Delta_P, X_R, X_I)}{2\Delta_P^2} X_P^2 \pi^T C_{PP} \pi \\
&\qquad + \frac{F(t, X_P, X_R+\Delta_R, X_I) + F(t, X_P, X_R-\Delta_R, X_I)}{2\Delta_R^2} X_R^2 C_{RR} \\
&\qquad + \frac{F(t, X_P, X_R, X_I+\Delta_I) + F(t, X_P, X_R, X_I-\Delta_I)}{2\Delta_I^2} X_I^2 C_{II} \\
&\qquad + \frac{X_P X_R \pi^T C_{PR}}{4\Delta_P \Delta_R}(F_k(t, X_P+\Delta_P, X_R+\Delta_R, X_I) - F_k(t, X_P-\Delta_P, X_R+\Delta_R, X_I) \\
&\qquad - F_k(t, X_P+\Delta_P, X_R-\Delta_R, X_I) + F_k(t, X_P-\Delta_P, X_R-\Delta_R, X_I)) \\
&\qquad + \frac{X_P X_I \pi^T C_{PI}}{4\Delta_P \Delta_I}(F_k(t, X_P+\Delta_P, X_R, X_I+\Delta_I) - F_k(t, X_P-\Delta_P, X_R, X_I+\Delta_I) \\
&\qquad - F_k(t, X_P+\Delta_P, X_R, X_I-\Delta_I) + F_k(t, X_P-\Delta_P, X_R, X_I-\Delta_I)) \\
&\qquad + \frac{X_R X_I C_{RI}}{4\Delta_R \Delta_I}(F_k(t, X_P, X_R+\Delta_R, X_I+\Delta_I) - F_k(t, X_P, X_R-\Delta_R, X_I+\Delta_I) \\
&\qquad - F_k(t, X_P, X_R+\Delta_R, X_I-\Delta_I) + F_k(t, X_P, X_R-\Delta_R, X_I-\Delta_I))
\end{aligned}
\tag{A.26}
$$

Policy improvement is performed pointwise for all $(x_1, x_2, x_3) \in \mathbf{P} \times \mathbf{R} \times \mathbf{I}$ using the following convex programming formulation:

$$
\begin{aligned}
\underset{\pi,\, \lambda_P,\, \lambda_R}{\text{minimize}} \quad & -\tilde{u}_k\left((1-\nu_I)(1-\lambda_P-\lambda_R)x_3\right) - \partial_P \hat{F}_k(\mathbf{c})\left((\mu - r\mathbb{1}_N)^T \pi x_1 + (1-\nu_I)\lambda_P x_3\right) \\
& - \partial_R \hat{F}_k(\mathbf{c})\left((1+\kappa)\lambda_R x_3\right) - \frac{1}{2}\partial_{PP} x_1^2 \pi^T C_{PP} \pi - \partial_{PR} x_1 x_2 \pi^T C_{PR} \\
& - \partial_{PI} \hat{F}_k(\mathbf{c}) x_1 x_3 \pi^T C_{PI} \\
\text{s.t.} \quad & \lambda_P + \lambda_R \leq 1, \\
& \pi^T \mathbb{1}_N \leq 1, \\
& \lambda_R \leq \gamma, \\
& \pi^T C_{PP} \pi \leq \sigma^2_{\max}, \\
& \pi, \lambda_P, \lambda_R \geq 0
\end{aligned}
\tag{A.27}
$$

BIBLIOGRAPHY

[1] Retirement Topics - IRA Contribution Limits — Internal Revenue Service. https://www.irs.gov/retirement-plans/plan-participant-employee/retirement-topics-ira-contribution-limits.

[2] Robo-Advisors - United States — Statista Market Forecast. https://www.statista.com/outlook/dmo/fintech/digital-investment/robo-advisors/united-states.

[3] Topic No. 409, Capital Gains and Losses — Internal Revenue Service. https://www.irs.gov/taxtopics/tc409.

[4] Abdullah Alshelahi. *Macroscopic Look at Equity Markets*. Thesis, 2019.

[5] Abdullah AlShelahi and Romesh Saigal. Insights into the macroscopic behavior of equity markets: Theory and application. *Physica A: Stat. Mech. Appl.*, 505:778–793, 2018.

[6] Andrew Ang and Joseph Chen. Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, 63(3):443–494, March 2002.

[7] Nikolaos Antonakakis, Juncal Cunado, George Filis, David Gabauer, and Fernando Perez de Gracia. Oil volatility, oil and gas firms and portfolio diversification. *Energy Economics*, 70:499–515, February 2018.

[8] Karina Arias-Calluari, Fernando Alonso-Marroquin, Morteza N. Najafi, and Michael Harré. Methods for forecasting the effect of exogenous risks on stock markets. *Physica A: Statistical Mechanics and its Applications*, 568:125587, April 2021.

[9] Francesco Barbieri, Jose Camacho-Collados, Luis Espinosa Anke, and Leonardo Neves. TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 1644–1650, Online, November 2020. Association for Computational Linguistics.

[10] Tessa Bauman, Bruno Gašperov, Stjepan Begušić, and Zvonko Kostanjčar. Deep Reinforcement Learning for Robust Goal-Based Wealth Management. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*, pages 69–80. Springer, 2023.

[11] Jason Baumgartner, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. The Pushshift Reddit Dataset, 2020.

[12] Maria P. Beccar-Varela, Maria C. Mariani, Osei K. Tweneboah, and Ionut Florescu. Analysis of the Lehman Brothers collapse and the Flash Crash event by applying wavelets methodologies. *Physica A: Statistical Mechanics and its Applications*, 474:162–171, May 2017.

[13] Christophette Blanchet-Scalliet, Nicole El Karoui, Monique Jeanblanc, and Lionel Martellini. Optimal investment decisions when time-horizon is uncertain. *Journal of Mathematical Economics*, 44(11):1100–1113, December 2008.

[14] Stephen P Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge university press, 2004.

[15] Sid Browne. Survival and Growth with a Liability: Optimal Portfolio Strategies in Continuous Time, July 1996.

[16] Jean L. P. Brunel. Goal-Based Wealth Management in Practice. *The Journal of Wealth Management*, 14(3):17–26, 2011.

[17] Jean L. P. Brunel. *Goals-Based Wealth Management: An Integrated and Practical Approach to Changing the Structure of Wealth Advisory Practices*. Wiley Finance Series. John Wiley & Sons, Inc, Hoboken, New Jersey, 2015.

[18] US Census Bureau. Income in the United States: 2022. https://www.census.gov/library/publications/2023/demo/p60-279.html.

[19] Yongyang Cai. Dynamic programming and its application in economics and finance. *PhD diss., Stanford University*, 2009.

[20] Yongyang Cai and Kenneth L. Judd. Shape-preserving dynamic programming. *Mathematical Methods of Operations Research*, 77(3):407–421, June 2013.

[21] Agostino Capponi and Yuchong Zhang. Goal Based Investment Management, July 2022.

[22] Anirban Chakraborti, Ioane Muni Toke, Marco Patriarca, and Frédéric Abergel. Econophysics review: I. Empirical facts. *Quant. Financ.*, 11(7):991–1012, 2011.

[23] Dwight Clancy. Average Fees for Financial Advisors in 2023. https://www.harnesswealth.com/articles/average-fees-for-financial-advisors/, May 2023.

[24] Julia Coronado and Karen Dynan. Changing retirement behavior in the wake of the financial crisis. 2011.

[25] John C Cox and Chi-fu Huang. Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory*, 49(1):33–83, October 1989.

[26] Sanjiv Das, Harry Markowitz, Jonathan Scheid, and Meir Statman. Portfolio Optimization with Mental Accounts. *The Journal of Financial and Quantitative Analysis*, 45(2):311–334, 2010.

[27] Sanjiv R Das and Daniel Ostrov. Dynamic Portfolio Allocation in Goals-Based Wealth Management. 2019.

[28] Sanjiv R. Das, Daniel Ostrov, Anand Radhakrishnan, and Deep Srivastav. Dynamic optimization for multi-goals wealth management. *Journal of Banking & Finance*, 140:106192, July 2022.

[29] Sanjiv R. Das and Subir Varma. Dynamic goals-based wealth management using reinforcement learning. *Journal Of Investment Management*, 18(2):1–20, 2020.

[30] M. H. A. Davis. Martingale methods in stochastic control. In M. Kohlmann and W. Vogel, editors, *Stochastic Control Theory and Stochastic Differential Systems*, volume 16, pages 85–117. Springer-Verlag, Berlin/Heidelberg, 1979.

[31] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, 2018.

[32] Francis X. Diebold and Kamil Yılmaz. Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, 28(1):57–66, January 2012.

[33] Francis X. Diebold and Kamil Yılmaz. On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1):119–134, September 2014.

[34] Matthew Dixon and Igor Halperin. G-Learner and GIRL: Goal Based Wealth Management with Reinforcement Learning, February 2020.

[35] Elizabeth Ashburn Duke. Changed Circumstances: The Impact of the Financial Crisis on the Economic Condition of Workers Near Retirement and of Business Owners : Remarks to the Virginia Association of Economists, Richmond, Va. March 2011.

[36] Mardi Dungey and Eric Renault. Identifying contagion. *Journal of Applied Econometrics*, 33(2):227–250, 2018.

[37] David Easley, Marcos M Lopez De Prado, and Maureen O'Hara. The microstructure of the "flash crash": Flow toxicity, liquidity crashes, and the probability of informed trading. *J. Portf. Manag.*, 37(2):118–128, 2011.

[38] Robert Engle. Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.

[39] H. Evensky, S.M. Horan, T.R. Robinson, and R. Ibbotson. *The New Wealth Management: The Financial Advisor's Guide to Managing and Investing Client Assets*. CFA Institute Investment Series. Wiley, 2011.

[40] Jason J Fichtner, John WR Phillips, and Barbara A Smith. Retirement behavior and the global financial crisis. *Reshaping Retirement Security: Lessons from the Global Financial Crisis*, pages 81–100, 2012.

[41] Jill E. Fisch, Marion Labouré, and John A. Turner. The Emergence of the Robo-advisor. (10), December 2018.

[42] Wolfgang Fruehwirt, Leonhard Hochfilzer, Leonard Weydemann, and Stephen Roberts. Cumulation, crash, coherency: A cryptocurrency bubble wavelet analysis. *Finance Research Letters*, 40:101668, May 2021.

[43] Petr Geraskin and Dean Fantazzini. Everything you always wanted to know about log-periodic power laws for bubble modeling but were afraid to ask. *The European Journal of Finance*, 19(5):366–391, May 2013.

[44] Kathrin Glau, Mirco Mahlstedt, and Christian Pötz. A New Approach for American Option Pricing: The Dynamic Chebyshev Method. *SIAM Journal on Scientific Computing*, 41(1):B153–B180, January 2019.

[45] Luigi Guiso, Paola Sapienza, and Luigi Zingales. Time varying risk aversion. *Journal of Financial Economics*, 128(3):403–421, June 2018.

[46] Nils H. Hakansson. Optimal Investment and Consumption Strategies Under Risk, an Uncertain Lifetime, and Insurance. *International Economic Review*, 10(3):443–466, 1969.

[47] Nils H. Hakansson. Optimal Entrepreneurial Decisions in a Completely Stochastic Environment. *Management Science*, 17(7):427–449, March 1971.

[48] Hamilton, James. Dates of U.S. recessions as inferred by GDP-based recession indicator. https://fred.stlouisfed.org/series/JHDUSRGDPBR, October 1967.

[49] Shu-Fan Hsieh, Chia-Ying Chan, and Ming-Chun Wang. Retail investor attention and herding behavior. *J. Empir. Financ.*, 59:109–132, 2020.

[50] Hugging Face. CardiffNLP/twitter-roberta-base-sentiment, 2022.

[51] C. Hutto and Eric Gilbert. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proc. Int. AAAI Conf. on Web and Soc. Media*, 8(1):216–225, May 2014.

[52] Daniel Kahneman and Amos Tversky. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263–291, 1979.

[53] Bong-Han Kim, Hyeongwoo Kim, and Bong-Soo Lee. Spillover effects of the U.S. financial crisis on financial markets in emerging Asian countries. *International Review of Economics & Finance*, 39:192–210, September 2015.

[54] Andrei Kirilenko, Albert S Kyle, Mehrdad Samadi, and Tugkan Tuzun. The flash crash: High-frequency trading in an electronic market. *J. Financ.*, 72(3):967–998, 2017.

[55] Rakesh Kochhar. A Recovery No Better than the Recession.

[56] Emanuel Kohlscheen, Benoît Mojon, and Daniel Rees. The Macroeconomic Spillover Effects of the Pandemic on the Global Economy, April 2020.

[57] Anoop S. Kumar and S. Anandarao. Volatility spillover in crypto-currency markets: Some evidences from GARCH and wavelet analysis. *Physica A: Statistical Mechanics and its Applications*, 524:448–458, June 2019.

[58] John V. Leahy and Joseph Zeira. The Timing of Purchases and Aggregate Fluctuations. *The Review of Economic Studies*, 72(4):1127–1151, October 2005.

[59] Shan Lei. Investment Strategies During the Great Recession: Who Remains Calm, and Who Panics?-The Role of Financial Planners. *Journal of Personal Finance*, 20(1):51–66, 2021.

[60] Alexander LeNail. NN-SVG: Publication-Ready Neural Network Architecture Schematics. *Journal of Open Source Software*, 4(33):747, January 2019.

[61] Yun Li. GameStop soars nearly 70%, trading briefly halted amid epic short squeeze, January 2021.

[62] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. RoBERTa: A Robustly Optimized BERT Pretraining Approach, 2019.

[63] Suwan(Cheng) Long, Brian M. Lucey, Larisa Yarovaya, and Ying Xie. "I Just Like the Stock": The Role of Reddit Sentiment in the GameStop Share Rally, September 2022.

[64] Štefan Lyócsa, Eduard Baumöhl, Tomáš Výrost, and Peter Molnár. Fear of the coronavirus and the stock markets. *Financ. Res. Lett.*, 36:101735, 2020.

[65] Anna Mancini, Antonio Desiderio, Riccardo Di Clemente, and Giulio Cimini. Self-induced consensus of Reddit users to characterise the GameStop short squeeze. *Scientific Reports*, 12(1):13780, August 2022.

[66] Harry Markowitz. Portfolio Selection. *The Journal of Finance*, 7(1):77–91, 1952.

[67] Walid Mensi, Shawkat Hammoudeh, Duc Khuong Nguyen, and Sang Hoon Kang. Global financial crisis and spillover effects among the U.S. and BRICS stock markets. *International Review of Economics & Finance*, 42:257–276, March 2016.

[68] Robert C. Merton. Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case. *The Review of Economics and Statistics*, 51(3):247–257, 1969.

[69] Robert C Merton. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3(4):373–413, December 1971.

[70] Edwin Ng, Zhishi Wang, Huigang Chen, Steve Yang, and Slawek Smyl. Orbit: Probabilistic Forecast with Exponential Smoothing, 2020.

[71] Efe A. Ok. *Real Analysis with Economic Applications*, volume 10. Princeton University Press, 2007.

[72] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zach DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. PyTorch: An Imperative Style, High-Performance Deep Learning Library, December 2019.

[73] Lasse Heje Pedersen. Game on: Social networks and markets. *Journal of Financial Economics*, 146(3):1097–1119, December 2022.

[74] Antonin Raffin, Ashley Hill, Adam Gleave, Anssi Kanervisto, Maximilian Ernestus, and Noah Dormann. Stable-baselines3: Reliable reinforcement learning implementations. *Journal of Machine Learning Research*, 22(268):1–8, 2021.

[75] Mikko Ranta. Contagion among major world markets: A wavelet approach. *International Journal of Managerial Finance*, 9(2):133–149, January 2013.

[76] Paul A. Samuelson. Lifetime Portfolio Selection By Dynamic Stochastic Programming. *The Review of Economics and Statistics*, 51(3):239–246, 1969.

[77] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal Policy Optimization Algorithms. https://arxiv.org/abs/1707.06347v2, July 2017.

[78] Parnia Shahrestani and Meysam Rafei. The impact of oil price shocks on Tehran Stock Exchange returns: Application of the Markov switching vector autoregressive models. *Resources Policy*, 65:101579, March 2020.

[79] T. A. Siddiqui, H. Ahmed, and M. Naushad. Diffusion of COVID-19 impact across selected stock markets: A wavelet coherency analysis. *Investment Management and Financial Innovations*, pages 202–214, 2020.

[80] Paolo Sironi. *FinTech Innovation: From Robo-Advisors to Goal Based Investing and Gamification*. John Wiley & Sons, 2016.

[81] Didier Sornette, Anders Johansen, and Jean-Philippe Bouchaud. Stock market crashes, Precursors and Replicas. *Journal de Physique I*, 6(1):167–175, January 1996.

[82] Pippa Stevens. AMC share price quadruples as retail traders raid hedge-fund short targets, January 2021.

[83] Lawrence H. Summers, David Wessel, and John David Murray. Rethinking the Fed's 2 percent inflation target. *Washington, DC: Brookings Institute*, 2018.

[84] Richard S. Sutton and Andrew Barto. *Reinforcement Learning: An Introduction*. Adaptive Computation and Machine Learning. The MIT Press, Cambridge, Massachusetts, nachdruck edition, 2014.

[85] Manolis N. Syllignakis and Georgios P. Kouretas. Dynamic correlation analysis of financial contagion: Evidence from the Central and Eastern European markets. *International Review of Economics & Finance*, 20(4):717–732, October 2011.

[86] Richard Thaler. Mental Accounting and Consumer Choice. *Marketing Science*, 4(3):199–214, 1985.

[87] Zaghum Umar, Afsheen Abrar, Adam Zaremba, Tamara Teplova, and Xuan Vinh Vo. The Return and Volatility Connectedness of NFT Segments and Media Coverage: Fresh Evidence Based on News About the COVID-19 Pandemic. *Finance Research Letters*, 49:103031, October 2022.

[88] Zaghum Umar, Mariya Gubareva, Imran Yousaf, and Shoaib Ali. A tale of company fundamentals vs sentiment driven pricing: The case of GameStop. *J. Behav. Exp. Financ.*, 30:100501, 2021.

[89] Evangelos Vasileiou. Does the short squeeze lead to market abnormality and antileverage effect? Evidence from the Gamestop case. *Journal of Economic Studies*, ahead-of-print(ahead-of-print), January 2021.

[90] Mark Veraar. The stochastic Fubini theorem revisited. *Stochastics*, 84(4):543–551, August 2012.

[91] Johannes Voit. *The Statistical Mechanics of Financial Markets*. Springer, Berlin, Heidelberg, 2001.

[92] John B. Walsh. An introduction to stochastic partial differential equations. In P. L. Hennequin, editor, *École d'Été de Probabilités de Saint Flour XIV - 1984*, pages 265–439, Berlin, Heidelberg, 1986. Springer Berlin Heidelberg.

[93] Charlie Wang and Ben Luo. Predicting $ gme stock price movement using sentiment from reddit r/wallstreetbets. In *Proceedings of the Third Workshop on Financial Technology and Natural Language Processing*, pages 22–30, 2021.

[94] Hungjen Wang, Anil Suri, David Laster, and Himanshu Almadi. Portfolio Selection in Goals-Based Wealth Management. *The Journal of Wealth Management*, 14(1):55–65,5, 2011.

[95] Eric W. Weisstein. Stone-Weierstrass Theorem. https://mathworld.wolfram.com/.

[96] Spencer Wheatley, Didier Sornette, Tobias Huber, Max Reppen, and Robert N. Gantner. Are Bitcoin bubbles predictable? Combining a generalized Metcalfe's Law and the Log-Periodic Power Law Singularity model. *Royal Society Open Science*, 6(6):180538, 2019.

[97] Jan Henrik Wosnitza and Cornelia Denz. Liquidity crisis detection: An application of log-periodic power law structures to default prediction. *Physica A: Statistical Mechanics and its Applications*, 392(17):3666–3681, September 2013.

[98] Dun-Zhong Xing, Hai-Feng Li, Jiang-Cheng Li, and Chao Long. Forecasting price of financial market crash via a new nonlinear potential GARCH model. *Physica A: Statistical Mechanics and its Applications*, 566:125649, March 2021.

[99] Yongan Xu, Jianqiong Wang, Zhonglu Chen, and Chao Liang. Sentiment indices and stock returns: Evidence from China. *Int. J. Financ. Econ.*, 2021.

[100] Menahem E. Yaari. Uncertain Lifetime, Life Insurance, and the Theory of the Consumer. *The Review of Economic Studies*, 32(2):137–150, 1965.

[101] Victor M. Yakovenko. Econophysics, Statistical Mechanics Approach to. pages 1–35. 2022.

[102] Xiaolan Yang, Yu Zhu, and Teng Yuan Cheng. How the individual investors took on big data: The effect of panic from the internet stock message boards on stock price crash. *Pac.-Basin Financ. J.*, 59:101245, 2020.